Parameter Search for Aesthetic Design and Composition

Henrik Carl-Olof Julian Oldfeldt Ekéus

Submitted for the degree of Doctor of Philosophy

Queen Mary, University of London

2016

I, Henrik Carl-Olof Julian Oldfeldt Ekéus, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third partys copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date: 23/06/2015

Publications

Ekeus H, Abdallah SA, Mcowan PW, Plumbley MD. (2013) 'How Predictable Do We Like Our Music? Eliciting Aesthetic Preferences With The Melody Triangle Mobile App'. Stockholm, Sweden: Proceedings of the Sound and Music Computing Conference 2013, SMC 2013, pp. 80-85.

Ekeus H, Plumbley MD, McOwan PW. (2012) 'Eye Tracking as Interface for the Design of Generative Visual Forms and Patterns'. Edinburgh, UK: Predicting Perceptions: Proceedings of the 3rd International Conference on Appearance, pp. 117-121.

Ekeus H, Abdallah SA, Plumbley MD, McOwan PW. (2012) 'The Melody Triangle: Exploring Pattern and Predictability in Music'. Musical Metacreation: Papers from the 2012 AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment. AAAI Technical Report WS-12-16, pp. 35-42.

Ekeus H, Mcowan PW, Plumbley MD. (2013) 'Eye Tracking as Interface for Parametric Design'. Paris, France : CHI 2013 Workshop on "Gaze Interaction in the Post-WIMP World"

Abdallah SA, Ekeus H, Foster P, Robertson A, Plumbley MD. (2012) 'Cognitive music modelling: An information dynamics approach'. 2012 3rd International Workshop on Cognitive Information Processing, CIP 2012

Collaborations

Chapter 3 presents some work carried out in collaboration. Two of the implementations of the *Melody Triangle* – the interactive installation (described in 3.4) and the desktop version (described in 3.5) – were developed as a collaboration between the present author and Samer Abdallah[1], as was the pilot study carried out with the desktop version of the triangle (described in 3.5.1). The interface to the mobile app version of the triangle was developed together with Tallah Ahmed at QApps[2]. All other aspects of the mobile app (the audio engine, the information dynamics engine, the social/crowdsourcing aspects), as well the server-side implementation of the crowdsourcing data collection was designed and developed by the present author. All analysis of the data collected by the mobile app is the work of the present author.

All remaining content of this thesis is the sole work of the present author.

[1]s.abdallah@ucl.ac.uk
[2]http://www.qappsonline.com/

Acknowledgements

I am enormously indebted to my supervisors Peter McOwan, Andrew McPherson, and Mark Plumbley. I would like to thank all of my family for there support. Additionally, special thanks go to all my colleagues on the MAT program. Further thanks go to Samer Abdallah, Richard Kelly, Pat Healey, Alexa van Sickle, Victoria Jones, Andy Clark, Joseph Sherwood, Barnaby Rowe, Martin Brooks, Pete Grant, Duncan Menzies, Anna Orhanen, Eduardo Cariglino, Antonio Arca, Nina French, Annemarie Gaillard, Jon Wallace, Steve Watson, Danny Sherwood, Conrad Molleson, Peter Neil, Bo Chen, Ana Crisan, Michael Tandy, Melanie Keeble, everyone at United Visual Artists, everyone at Heresy AI, and everyone else who has supported me over these years.

This work was funded by the Engineering and Physical Sciences Research Council (EPSRC) as part of the Doctoral Training Centre in Media and Arts Technology at Queen Mary University of London (ref: EP/G03723X/1).

Parameter Search for Aesthetic Design and Composition

Henrik Carl-Olof Julian Oldfeldt Ekéus

Abstract

This thesis is about algorithmic creation in the arts – where an artist, designer or composer uses a formal generative process to assist in crafting forms and patterns – and approaches to finding effective input parameter values to these generative processes for aesthetic ends.

Framed in three practical studies, approaches to navigating the aesthetic possibilities of generative processes in sound and visuals are presented, and strategies for eliciting the preferences of the consumers of the generated output are explored.

The first study presents a musical interface that enables navigation of the possibilities of a stochastic generative process with respect to measures of subjective predictability. Through a mobile phone version of the application, aesthetic preferences are crowd-sourced.

The second study presents an eye-tracking based framework for the exploration of the possibilities afforded by generative designs; the interaction between the viewers' gaze patterns and the system engendering a fluid navigation of the state-space of the visual forms.

The third study presents a crowd-sourced interactive evolutionary system, where populations of abstract colour images are shaped by thousands of preference selections from users worldwide

For each study, the results of analyses eliciting the attributes of the generated outputs – and their associated parameter values – that are most preferred by the consumers/users of these systems are presented.

Placed in a historical and theoretical context, a refined perspective on the complex interrelationships between generative processes, input parameters and perceived aesthetic value is presented.

Contributions to knowledge include identified trends in objective aesthetic preferences in colour combinations and their arrangements, theoretical insights relating perceptual mechanisms to generative system design and analysis, strategies for effectively leveraging evolutionary computation in an empirical aesthetic context, and a novel eye-tracking based framework for the exploration of visual generative designs.

Submitted for the degree of Doctor of Philosophy

Queen Mary, University of London

2016

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The use of generative processes for the discovery of novel forms and patterns is not uncommon to the practices of artists, composers and designers. In western music there exists a long tradition of algorithmic thought; from the isorhythms and canons of early music, through serialism to the controlled 'chance music' of John Cage and the 'formalised music' of Xenakis. Similarly from the religious art of the Islamic world to conceptual visual art pieces of the 20th century there are numerous examples of the application of abstract systems to the creative process.

In recent decades, however, the use of generative processes has accelerated due to the increased availability of computing power. As illustrated by the ubiquity of computer based digital art forms, code has become an artistic medium in its own right. In some cases the used systems are simply tools to an end, in other cases the algorithm, the process itself, becomes the object of artistic inquiry.

A fundamental motivation of this research is to gain a better understanding of algorithmic creation. Generative process are often found in both music and art across cultures, but what is a generative process? What is algorithmic creation? And what is it that compels artists and composers, consciously or not, to so often employ these techniques? These are the broad questions on which this research seeks to shed but a little light.

At its simplest, algorithmic creation can be understood as consisting of a formal *generative system* or process, that takes *parameters* as input, and that generates some sort of $output$. With this view, the composer/designer's task in algorithmic creation can be broken down into two steps; that of creating/formulating the generative process or algorithm, and that of finding the appropriate parameters to input into the system. This research is primarily focused on issues surrounding this second step of *parameter search and discovery*.

Algorithmic thought and practice afford artists and composers new creative possibilities; they make attainable forms and patterns that would be difficult to reach otherwise, while also opening for the possibility of automating aspects of art and music making. However they also come bundled with difficulties and peculiarities; generative designs are difficult to control, and difficult to predict. In particular, in this mode of artistic creation there can be numerous intermediary layers between an artist's action and its consequences. This brings up difficult questions for the designers of creative algorithmic systems, and an aim of this research is to explore frameworks and interfaces to facilitate the process of parameter search and discovery.

This research employes generative systems as a framework through which questions of aesthetics value, taste and subjectivity are raised and explored. While the broad questions of aesthetics have been the subject of philosophical enquiry for centuries, this research re-visits some of these questions in the light of algorithmic creation and through the prism of parameter search and discovery.

Parameter selection is a more constrained and controllable process than unbounded art or music making, and better lends itself to methods of scientific analysis. However choosing parameters for an artistic or musical generative system nonetheless involves making a judgment of value on the generated output. By collecting parameter selections from people interfacing with artistic or musical generative systems, there comes an opportunity to be able to uncover, measure and compare the tastes of individuals. This is the main approach taken in this thesis; three practical studies elicit parameter selections from individuals, and these are then analysed to look for trends in aesthetic preferences.

Looking for these trends can help shed light on the slippery issue of aesthetic objectivity and subjectivity. Is it true that there is 'no accounting for taste', or is there a normative dimension to aesthetic judgments? By averaging across the judgments of multiple individuals, it may be possible to find and measure what is common to all of them. The analyses of the practical studies in this thesis all try to seek out this *average*, to filter through idiosyncrasies of taste and get at underlying commonalities; to reveal the *objective* dimension of aesthetic judgments. This can yield insights into the underlying cognitive and perceptual mechanisms at play, while bringing new possibilities for the automation of music and art creation and the design of smart creative interfaces.

1.1 Aims of Research

The aims of this research can be summarised as follows:

- To elucidate the nature of algorithmic creation, and the relationship between generative processes, parameter search, and aesthetic value.
- To explore novel mechanisms and interfaces for parameter discovery and the exploration of the aesthetic possibilities of generative processes.
- To discover cross-user trends in aesthetic preferences in visual patterns and musical preferences.

1.2 Approach

This thesis begins with a historical review of generative processes in art and music, identifying common themes and theoretical concepts. This is then followed by an exposition on the theoretical nature of generative systems, bringing together ideas from information theory, cybernetics and cognitive science. This forms the theoretical grounding and language through which the research is carried out and evaluated, and is the subject of chapter 2.

The three practical explorations of this research each embody contrastive approaches and methods to this question: *given all the possibilities afforded by a generative process, how can we find the input parameters that will yield desirable results?* Each of these practical studies concern different modalities; one is musical, one generates abstract black and white patterns, the third generates patterns of colour.

The first of these practical explorations is the *Melody Triangle*; a musical interface that enables users to explore melodies generated by stochastic processes, and that as a mobile phone application collects the preferred settings of its users. The *Melody Triangle* is the subject of chapter 3. The second practical exploration, *Keyebernates*, enables a slow semi-guided navigation of the possibilities afforded by a visual generative process. This is done by balancing randomising noise processes against a volition of the user, as interpreted from gaze data gathered in real time through an eye-tracking interface. *Keyebernates* is the subject of chapter 4. The third is *EvoColour*, a crowdsourced interactive evolutionary system that generates simple coloured images, selections from users in the wild driving the genetic algorithm, the subject of chapter 5.

Although the three studies seem very different, as they concern differing sensory modalities with differing control mechanisms, they share a core equivalence: the three systems are frameworks that 'wrap round' fundamentally simple generative processes at their core (stochastic melodies for the *Melody Triangle*, patterns of intersecting circles for *Keyebernates*, and coloured circular images for *EvoColour*) and provide an alternative to manual parameter search for these processes. They are systems for parameter search and discovery. The abstract nature of generative processes is such that they can be studied independently of the medium in which they may take their final form. Tackling differing domains together in one body of research, although challenging, provides the opportunity for uncovering the common currents underlying algorithmic creation in the arts.

The practical studies split the artistic decision making processes across multiple actors. The generative systems are designed as part of the setup of the experiments, while the task of parameter selection, and of evaluation of the aesthetic worth of the works, are assigned to the participants of the studies. This raises issues of creative responsibility and attribution that are difficult to solve, and will be re-visited throughout this thesis.

Each of the systems in these studies collect feedback from users in some way, and analyses are carried out on this data to extract the features and attributes that are deemed 'desirable' by users of the systems, as well as those that are 'undesirable'; namely colour combination preferences for *EvoColour*, musical attributes for the *Melody Triangle* and visual patterns for *Keyebernates*. This user feedback, although mediated through different interfaces, ultimately distills to sets of parameter values for the core generative processes at the centre of each of these systems.

The design of the experiments aims to establish a correlation between the parameter values selected and the participants' judgments of the aesthetic value of the generated output. With this collected sets of parameter values in hand, the analyses of each of the experiments employ analytical and numerical techniques to search for trends in this data. This is done by looking for averages and peaks of popularity across the users' inputs and the features of the generated outputs. If aesthetic judgments have both a common objective dimension and a subjective dimension, then it is hoped that these averages are a way to identify the objective, common elements of the aesthetic responses; the recurring patterns across individual tastes.

Finding these patterns can yield insights into mechanisms of perception. Additionally, these insights may be used to *predict* what kinds of outputs may be deemed of aesthetic value. This can have numerous practical applications, particularly when it comes to the design of systems for the automatic creation of artistic artefacts.

Naturally there are a number of caveats. For instance, how well do these identified trends the systems extract from the users *really* correspond to aesthetic value? This is dependent on wether or not the experimental setup successfully aligns user inputs with aesthetic preference. Further one must be certain that identified patterns are not due to some other factors, such as interface priming.

1.2.1 Success

How then is one to know if these studies has been 'successful'? Here a criteria for success is proposed. This criteria is then used in the subsequent chapters to evaluate each of the three practical studies in turn.

Success consists of meeting the following two *criteria*:

C1: *correlation between parameter choices and aesthetic value of the output*. A key criteria is that the user's selection of parameter values must align with their aesthetic judgment of the generated output. Did the participants really select a set of parameters due to the aesthetic properties of the generated artefact? Or have they been selected for other reasons, such as biases introduce by interface priming?

C2: *sufficient volume of data points and significant trends*. The space of all possible input parameters for even relatively simple generative processes can be large and high dimensional. A key criteria for success is wether or not there are enough input data points, relative to the size of the parameter space, to sensibly suggest that there is a statistical significance to the trends and averages found. Further as some of these studies are crowd sourced, individual participants will have been subject to differing environmental conditions as well as presentation conditions. As such there would need to be enough data points to average out these effects.

1.3 Contributions of This Thesis

The contributions of this thesis can be divided into two areas: the identifications of trends in aesthetic preferences from the collected data, and the three practical explorations as epistemic artefacts in their own right, in particular, the theoretical and practical insights gained in the development of these systems and the running of the experiments.

1.3.1 Trends in Aesthetic Preferences

EvoColour

Of the three, *EvoColour* was the framework that was most successful in identifying patterns of aesthetic preference across users, and the only experiment that could be said to match the criteria for 'success' defined above in section 1.2.1. In *EvoColour*, participants repeatedly indicate their aesthetic preference between two images composed of concentric circles of up to three colours, called *Markov Images*[1]. The cumulative judgments drive a 'survival of the fittest' process whereby the populations of images evolve in response to the aesthetic judgments of the users.

Analyses of the many thousands of selections, over generations of evolving images, yielded numerous insights. These included identifying trends in preference with regards to overall colour, colour combinations, as well as the arrangements and relative amounts of theses colours in an image. The observations that were made for *EvoColour* are detailed in section 5.7.

Some of these observations support previous work in colour research. For instance the well known observations that, overall, blue is the 'world's favourite colour', and that yellows are disliked, were echoed in the analysis of the data collected by *EvoColour*. Additionally some less well known patterns in colour combination preferences were also reflected, for instance the observation that high contrast in lightness in an image is generally deemed desirable.

However many new observations of aesthetic preference, not previously identified in the literature of colour research, were brought to light with *EvoColour*, in particular, observations relating to the *arrangements* of colours in an image, as well as the *relative amounts* of these colours. This includes the observations that both images that embody perfect repetition and those that lack all structure were disfavoured over images of intermediate regularity, that sparsity is favourable, and that images with unequal amounts of colour are preferred. Additionally it was

[1]Named after the Markov process that determines the sequence of the colours. This is explained in detail in section 5.4.1

shown that colour preferences are highly contingent on how these colours are arranged (images of high entropy had differing overall colour preferences to images of low entropy); an area not previously explored in the literature.

Melody Triangle, Keyebernates

In the other two studies, the *Melody Triangle* and *Keyebernates*, the findings with regards to identified trends in aesthetic preference were modest. Neither of these two studies matched the criteria for success defined in 1.2.1.

With the Melody Triangle mobile app, aesthetic preferences in musical patterns were elicited by collecting the most popular settings from users in the wild. As oulined in section 3.7, it was identified how there was a significant dislike for unpredictability and a strong preference for predictable musical patterns. Given music's fundamentally patterned nature however, this is perhaps not so surprising. Other potentially insightful observations were likely tainted by bias introduced through interface priming (failing success criteria $C1$). Additionally this study was hampered by difficulty in gathering a large enough user-base relative to the size of the parameter-space (failing success criteria $C2$).

In *Keyebernates*, the possibilities of a simple generative mechanism (consisting of overlapping circles) are explored through a gaze-tracking interface, the parameter values most gazed upon by the users being reinforced. In the analysis in section 4.4 it was shown that there is a trend towards images of greater complexity, as all users pulled the system towards circles of increased size. However it was not possible to determine if this was due to an aesthetic preference for these patterns that gaze was drawn to them (failing success criteria $C1$). Further the system induced fatigue in the participants, this combined by an overall slowness in parameter space navigation made it difficult to gather enough data points (failing success criteria $C2$).

As such the *Melody Triangle* and *Keyebernates* were not particularly successful in yielding novel insights into aesthetic preferences.

1.3.2 Epistemic Artefacts and Theoretical Insights

The three practical studies of this thesis were not only created as a means of finding trends in aesthetic preferences or 'the best' parameter values for any particular generative process. The *Melody Triangle*, *EvoColour* and *Keyebernates* can also be seen as *epistemic* artefacts; the process of designing and constructing these systems, and analysing their gathered data, provided avenues through which theoretical insight into the inter-relationships between generative process, parameters and perceived aesthetic value could be gleaned.

Conceptual Spaces as a Framework for Generative Processes

One such insight is that understandings of the mechanisms of perception can greatly benefit the design of tools and interfaces for generative processes. More specifically, a contribution of this thesis is to present the applicability Peter Gärdenfors *Theory of Conceptual Spaces*(Gärdenfors, 2004b) to the design of generative systems, the development of interfaces, and to analysis.

The theory of Conceptual Spaces, presented in section 2.4.2 is a framework for representing concepts and describing objects based on geometrical structures. Within conceptual spaces, similarity is related to proximity in a geometrical space defined by a set of *quality dimensions*, each representing a subjective 'quality' of the object.

Generative processes often have a difficult to predict interrelation between parameters and generated output. This thesis suggests that an effective mechanism for control can be attained if parameters are aligned to the relevant quality dimensions of the perceived output.

This is demonstrated with the *Melody Triangle*, where the insights of *Information Dynamics*, an information theoretic model of subjective predictability, presented in section 2.6.4, helps uncover some of the underlying conceptual spaces of (an aspect of) the perception of music. This knowledge was successfully applied to the design of the *Melody Triangle* interface, where the control parameters align with the quality dimensions of predictability. The *Melody Triangle* provides a new representation, a particular way of conceiving and understanding sequences of temporal events. It determines the descriptions by which a signal might be acted upon, contemplated and understood, thus orientating possible modes of interaction with respect to that signal that could not be achieved otherwise.

Similarly for *EvoColour*, it is shown how an understanding of the relevant conceptual space not only affords better ways to analyse the collected data and to reason about colour, but in the implementation of its algorithms, demonstrated how it could form an effective basis for the design of a creative system.

Interactive Evolution in Empirical Aesthetics

Another contribution of this thesis is the methodology that *EvoColour* embodies. Interactive evolution has commonly been used to generate aesthetic artefacts. However these generative processes often have intangibly large state-spaces, and their generated artefacts complex and difficult to measure features. This makes it difficult to draw significant findings in an empirical aesthetics context. In *EvoColour* the state-space of the evolving *Markov Images* are much smaller that in many other evolutionary systems, and relevant features are easily identified. In the analysis for *EvoColour*, it was shown how many more observations could be drawn from the evolving populations of images, than a non-evolving 'control' population. This suggests that the approach taken – using interactive evolution on a generative system with a state-space of modest size and easily extractable features – is a fruitful approach for empirical aesthetics.

Gaze as Control Mechanism

Keyebernates, as an instance of scientific and artistic exploration, considers the act of perceiving, of looking itself, as an input modality for algorithmic systems. A contribution of this thesis was *Keyebernates*' novel mechanism by which volitional-gaze could be used to navigate the parameter-space of generative designs. This is done through a delicate balancing and tuning of a normalising force driven by gaze, against a disturbing force driven by noise processes, this interaction enabling a slow semi-guided navigation of parameter space. It was shown that when participants are asked to look for certain patterns, that they can guide the system towards the corresponding area of parameter space. However the navigation was slow, and further work is needed to make *Keyebernates* practical in a design context.

1.4 Thesis Structure

Chapter 2 provides the background to this research. After stating the overall context of research, a historical survey of generative process in the arts is provided in section 2.2. Section 2.2.1 begins with a discussion on the nature of formalisms in music. This is followed by a historical survey of generative music before computers. An overview of generative music with computers is then provided in section 2.2.2.

Looping back in time, the use of formal and algorithmic processes in the history of the visual arts are discussed in section 2.2.3. This leads on to a section on computer art in section 2.2.4, covering the earliest experiments straight through to modern uses of evolutionary computation in art and music.

Section 2.3 presents an exposition on the theoretical nature of generative processes. Concrete examples, formal definitions and key concepts are provided. This chapter includes a discussion on the relationship between generative processes and constraints(section 2.3.1) parameterisation (section 2.3.2) and stochastic processes (section 2.3.3).

Section 2.4 continues the exposition and discusses issues related to the perception or consumption of the artefacts of generative systems. It brings in subjective notions of similarity (section 2.4.1) and discusses how they can be measured, and these are linked to Gärdenfors' *Theory of Conceptual Spaces*(Gärdenfors, 2004b) in section 2.4.2.

Section 2.5 provides a discussion on the nature of self-organisation, discusses how it may be quantified (section 2.5.1) and its link to trend identification and data collection (section 2.5.2).

Section 2.6 concerns aesthetics. Some philosophical issues are briefly discussed and linked to generative art in section 2.6.1. An overview of empirical aesthetics is provided in section 2.6.2, and efforts to measure aesthetic value are discussed in 2.6.3. Section 2.6.4 discusses information theory and aesthetics, and presents *information dynamics*, the context for the first practical study of this research, the subject of the following chapter. Finally section 2.6.5 discusses issues of computational aesthetics.

Chapter 3 concerns the *Melody Triangle*, a musical interface for the exploration of stochastically generated content. Section 3.2 describes the information theoretic measures behind the triangle. How these are then applied to the construction of the triangle is described in section 3.3. Section 3.4 describes the first incarnation of the triangle: the interactive installation. Section 3.5 describes the second incarnation of the triangle – the desktop application – as well as a pilot study carried out with it. Section 3.6 concerns the *Melody Triangle* mobile phone application. Its features are outlined, including how it crowd-sources user preferences, and the results of the analysis are presented. A discussion of the results and avenues of potential further work are outlined in section 3.7.

Chapter 4 concerns *Keyebernates*, an eye-tracking based system for the navigation of generative designs. Section 4.1 sets the context, with a discussion of previous work on gaze and interest in section 4.1.1, and an overview of eye-tracking based applications in section 4.1.2. The design of *Keyebernates* is outlined in 4.2, and a study carried out with it is described in 4.3. In section 4.4, the results of the study are discussed, avenues of potential further work are outlined, and the significance of *Keyebernates* as an aesthetic object in its own right is put forward.

Chapter 5 concerns *EvoColour*, a crowd-sourced interactive evolutionary algorithm for the evolution of coloured images. Background on colour theory is provided in section 5.2, and a brief survey of the literature on colour preferences is provided in 5.3. The design of *EvoColour* is outlined in 5.4, and a detailed analysis of the collected data is provided in 5.6. Section 5.7 provides an overview of the numerous observations derived from the collected data, and avenues of further research are outlined.

Finally chapter 6 revisits the three practical studies and discusses their significance in a broader theoretical context.

Chapter 2

Background

2.1 Context of Research

Choices and decisions are a natural part of the artistic, design and composition process. A graphic designer chooses fonts or colours, an artist chooses forms, a composer picks and arranges notes, melodies and chords, and the writer generates meanings by continually choosing words from the language at every step. All canvases, in all media, represent a space of possibilities, and a specific 'work' or artefact can be viewed as one of the possibilities in this space, brought about by choices made by the artist.

The creative processes can be mediated by sets or rules that reduce the number of possible choices. These rules can take the form of social conventions, such as those defined by 'style', 'genre' or language. These rules can be strict, or socially defined norms that can be bent. Familiarity with the methods and procedures of styles and genres is part of the artist's craft, and the rules that define styles and genres themselves are fluid and evolve over time, and form a lens through which works are evaluated by the wider community(Boden, 2004).

The reduction of the space of possibilities can be great; the rules of Haiku poetry for instance place restrictions on both thematic and structural arrangement of syllables. Similarly the mid-twentieth century composition style of 'serialism' places great restrictions on the composers' choices of notes in a melody. However even in these cases, the space of possibilities is still vast.

In other cases the composer or designer consciously creates the set of rules that reduce the space of possible choices. This is the practice of 'algorithmic composition' or 'generative design'. Here the set of rules can simultaneously be seen as a set of restrictions that define a set of criteria on the output, or the rules become a 'generative process' that take in a set of inputs and *generate* an output. In using such generative procedures, new and unexpected forms that would be difficult to otherwise conceive become available. Simultaneously all the forms that cannot be created through the generative process disappear from the space of possible outputs. This distinction and relationship between rules and stylistic conventions is subtle.

Even composers and designers who engage with generative processes still must make decisions, often in the form of *parameter search*, one of the primary concerns of this research.

Algorithmic composers and designers – those that *explicitly* engage with generative processes – have become increasingly prevalent with the ubiquity of computers, so much so that computational substrate has come to be considered an artistic medium itself, as evident with the emergence of 'computer art' and 'software art'. However as will be outlined in the next section, this is by no means a process new to the computer era.

2.2 Historical Foundations

It this section, an extensive survey of the use of generative processes in the arts is presented. Key historical examples are discussed and recurring themes and undercurrents are highlighted.

2.2.1 Music

The Codifiability of Music

Music and mathematics are intertwined in the most fundamental way; mathematical concepts such as ratios and sets are intrinsic to the understanding of scales, harmony and rhythm. From the earliest Pythagorean explorations of string subdivisions, to modern 'live-coding' laptop performances, formalised procedures and structures are pervasive to the history of music.

With regards to western tradition, one need only consider a musical score to see ample evidence; pitch is organised in sets - dividing frequencies into scales, time is divided and subdivided in metric structures. The composer and sound artist Trevor Wishart describes this quantisation as *lattice sonics*(Wishart & Emmerson, 1996, p. 8); the grids that divide time and pitch, around which musical styles and conventions have evolved. The lattice sonics, embodied in musical notation, makes music inherently *codifiable*, and hence suited to formalised structures and procedural methods. As Rothgeb suggests, "the printed score and its representation in an appropriate linear code, incompleteness notwithstanding, present a sufficiently detailed picture of the music's structure to render it accessible to analytical probing in a depth not approachable in many other art-forms"(Rothgeb, 1980).

The forms of western music have been codified as pattern descriptions on this lattice; rondo, isorhythms, canon, sonata all examples of structural frameworks for musical composition. Wishart notes (and laments) how this codifiability and formalism is reflected in the syllabi of western universities, where "tremendous emphasis [is] placed on the study of composers who employ a clearly, rationally, codifiable (verbalisable), musical praxis, in particular the work of Palestrina .. J.S. Bach and of course, Schoenberg and his 'twelve-tone technique"'(Wishart & Emmerson, 1996, p. 15).

Such codified descriptions bare an intimate relationship to generative processes; for example a *cannon* can be considered not only a structural form, but also a generative process, as it defines a rigid set of constraints on note and duration selections in the lattice. In a cannon, multiple, identical, but transposed melodies, are played after one another. As the notes of one melody are put to paper, the constraints of cannon form effectively cause the other melodies to be 'generated'. As such, the formalised description of the cannon can be understood as an algorithm and the starting melody as the input 'parameters': constraints in form and generative processes are interlinked.

It it not always clear when to call a musical work 'generative'; if all the notes of a composition are the result of a formal generative procedure, then it would be fair to consider it a generative composition. But is that still the case if, for example, half the musical content was generated in this way? If the formal procedures are entirely 'in the head' of a composer, as the internalised conventions and norms in which she operates, is this still a 'generative' process? Or are they merely 'constrained' to the formalisms of style?

Loy distinguishes between 'structural form' and 'procedural form' of a musical work: "Structural form refers to its analytic deep structure and the syntactic rules relating to it.. The procedural form refers to the methods governing its actual fabrication and composition process...All formal systems, wether they are procedural or structural, are mechanisms of externalizing of at least some of the composer's decision making on to some automaton of some sort"(P. Todd & Loy, 2003). This 'externalizing' enables the composer to explore musical material that would otherwise lie out of reach; the constraints demarcate the *design space* within which the composer operates.

Building on the long history and tradition of style, structural forms help frame and guide the composer's output (a composer born on a desert island is unlikely to compose a fugue). Similarly a procedural form can frame and guide a musical composition within a generative, parameterised context. However as Vaggione articulates, "the role of the composer here is not one of setting a mechanism and watching it run, but one of setting the conditions that will allow him or her to perform musical actions", for composers are engaged in a *performative* act with "musical materials emerging concretely out of a critical interaction with their materials, including their algorithms.. formalisation is not foundational, but operational, local, and tactical"(Vaggione, 2001).

Local to each composition or work, the composer chooses, builds and continually adjust the formalism, acting on the results. Formalism operates *in situ*, within the artistic process: constraints are created, their generative consequences continually perceived and acted upon as part of a perception/action cycle. Structural forms *incidentally* facilitate the composer's task, in so far as they relieve the composer of choices to be made, whereas generative procedures do so *explicitly*. The degrees to which the formalisms are explicitly articulated vary across genres, composers, and across each individual musical work. The choice of appropriate terminology, wether to call a work 'generative', is a reflection on how visible and externalised the formalisms are, rather than a binary property of a particular artefact.

A Very Brief History

The history of Western religious and classical music followed a trajectory of increasingly intricate patterns of *lattice sonics*, from Gregorian chant and isorhythms, the rules of counterpoint as embodied in Johann Joseph Fux's *Gradus ad Panassum*(Fux, Mann, & Edmunds, 1965), to the crystalline fugues of J.S. Bach. As romanticism brought emotion to the fore, the mathematical structures became less explicit, yet were still present as the conventions that Schoenberg, (with the advent of serialism), sought to break free from. A liberation of sorts, but only perhaps to replace it with even more explicit and strict processes and procedures as embodied in the 12-tone technique. Xenakis summarised this moment in music history: "atonal music broke up the tonal function and opened up a new path parallel to that of the physical sciences, but at the same time constricted by the virtually absolute determinism of serial music"(Xenakis, 1992, p. 4). This new alternate formalism pre-occupied the avant-guard, and was itself evolved and iterated upon by students of Schoenberg, such as Webern, Berg, and Boulez, and influenced many others such as Stockhausen, Messiaen and Xenakis.

Schoenberg's great achievement was recognising the possibility of a music not bound by the tonality-based formalisms of classical tradition; he understood tradition as a set formal structures, and as such that they could be replaced. However basing a composition on rigorous formal logical constructs does not guarantee a musically coherent result, as there is no certainty that a listener can *perceive* the internal logic of the structures. Serialism itself would not readily yield music widely favoured outside the bubble of the avant-guard, however the break with tonality opened the floodgates; composers would now freely define their own formal structures.

Some music is more readily perceived as procedural or algorithmic then others. The distinction between structural form and generative form was muddied in the highly pattern music of minimalism, as exemplified by composers such as Steve Reich, Terry Riley, Philip Glass and Arvo Pärt. Here, the geometric structures of patterns are readily heard and quite easily perceived by the listener. Steve Reich articulates well this confluence of formal processes with perceived form when referring to his 'Music as Gradual Process': "I am interested in perceptible processes. I want to be able to hear the process happening throughout the sounding music .. What I'm interested in is a compositional process and a sounding music that are one and the same thing"(Reich, 2004, p. 35).

Non-western music has not yet been mentioned here, and space limits considerations, however one need only briefly consider the looping structures of Gamelan and the many traditional African phasing rhythmic patterns, to see that procedural elements are not limited to western music.

Indeterminacy

John Cage, Schoenberg's most famous pupil, brought in indeterminacy explicitly into his music[1], leaving large swaths of the composition determined by chance operations.This employment of controlled chance would leave elements of the performance underspecified, many of the details to be either improvised or chosen by performers (this is known as 'aleatoric music', a technique also approached by composers such as Stockhausen). Chance would also be applied to the earlier step of defining the composition, such as in Cage's *Music of Changes* (1951) a work where he attempted to define all musical parameters by chance processes(Nyman, 1999, p. 6).

[1]Although he was not the first, aeolian harps and wind chimes ancient examples.

Despite this indeterminacy, aleatoric music can be seen as generative; high-level constraints determine the general course, but the details are generated by chance operations. Cage's influence resonates still today, with the work of composer/producers such as Brian Eno. He articulates well the generative aspect of indeterminacy - "Ordinary music is like engineering, where everything's built according to a plan, and it's the same every time you play it. Generative music is more like gardening; you plant a seed, and it grows different every time you plant"(as quoted in (Essl, 2007, p. 122)).

Chances operations were also approached scientifically. Advances in information and communication theory, as embodied by Shannon and Weaver's *Mathematical Theory of Communication*(Shannon & Weaver, 1964), inspired a number of attempts at musical composition using stochastic processes. Shannon and Weaver explicitly suggest music is a domain suited to their approach, stating that communication systems "include all of the procedures by which one mind may affect another. This, of course, involves not only written and oral speech, but also music, the pictorial arts, the theatre, the ballet, and in fact all human behaviour"(Shannon & Weaver, 1964, p. 3). The authors use Markov chains to model the probabilities of message formations as symbols, which they define as follows:

> "A system which produces a sequence of symbols (which may, of course, be letters or musical notes, say, rather than words) according to certain probabilities is called a stochastic process, and the special case of a stochastic process in which the probabilities depend on the previous events, is called a Markoff process or a Markoff chain"(Shannon & Weaver, 1964, p. 11).

With Markov processes, Shannon and Weaver could generate new sentences from English letter and word transition probabilities(Shannon & Weaver, 1964, p. 43).

It would not be long before they inspired a number of pre-computer experiments that applied Markov processes to the generation of musical content. In those experiments, the selections of musical materials are acquired through coin-flips, dice or weighted decks of cards, the Markov chains reflecting statistically simple melodic sources such as nursery rhythms or cowboy songs(Ariza, 2005b, p. 41). Markov processes however are quite tedious to run manually, and with pseudo-random number generators, computers would become a natural medium for such experiments.

2.2.2 Generative music with computers

Given that procedural thought permeated music practices even before the advent of the computer, it would be a natural progression for music to be explicitly applied to computational and algorithmic processes. As such, when computers make their appearance in academia and research centres, it would not take long for the first experiments in algorithmic music composition and sound synthesis to occur. As general symbolic processors, computers naturally engender an emphasis on the formal aspects of musical activity.

Long before then, in 1843, Ada Lovelace famously articulated music's potential for explicit automation. When commenting on Babages' (never built) Analytical Engine, she suggests that "it could act upon other things besides number". Suggesting that "if the fundamental relations of pitched sounds in the science of harmony and of musical compositions" could be appropriately symbolically represented as input, that "the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent" (McHale, 2011). The fact that she attributes the act of composition to the engine, and not the programmer of the engine, foreshadows future debates on attribution and authorship in algorithmic creation.

Research into the application of computers to music and composition began in the 1950s, and continues to this day. The first experiments in computer-based music creation occurred in 1951 with the simple synthesised melodies programmed by Geoff Hill on the CSIR MK1 computer(Doornbusch, 2002), closely followed Bolitho and Klein's early experiments in composition, where randomly selected pitches were filtered with constraints encoding musical rules(Hiller & Isaacson, 1959, p. 55). A more famous early example is the *Illiac Suite* of Hiller and Isaacson(Hiller & Isaacson, 1959, p. 6). A series of movements for string quartets were composed with the Illiac computer, each movement embodying an encoding of different stylistic constraints(Ariza, 2005b, p. 64). The *Illiac Suite* used stochastic methods including Markov processes, encoded formal rules based on Fux's rules of counterpoint, as well as encoding 20th century serialist procedures. Space considerations prevent a detailed look at this seminal work, but a detailed description of each of the movements, called *Experiments*, can be found in (Ariza, 2005b).

Iannis Xenakis famously said "music, by its very abstract nature, is the first of the arts to have attempted the conciliation of artistic creation with scientific thought. Its industrialization is inevitable and irreversible"(Xenakis, 1992, p. 123). He goes on to suggests the *Illiac Suite* as concrete evidence of this motion: "Have we not already seen attempts to industrialize serial and popular music.. by the musicological research of Hiller and Isaacson?"(Xenakis, 1992, p. 123). However novel the creation of the *Illiac Suite*, its basis on musical idioms made the resulting works reflect existing musical styles (hence Xenakis referring to it as 'musicological research').

Xenaxis describes his approach to music in his text *Formalized Music*(Xenakis, 1992) with little emphasis on modelling of existing idioms and forms, but states instead "the first task is to construct an abstraction from all inherited conventions and to exercise a fundamental critique of acts of thought and their materialization"(Xenakis, 1992, p. 11). He pioneered a wide gamut of original procedural techniques and applied them to the creation of new and never attempted sound worlds, both with and without computers. In particular he brings in a scientifically informed non-determinism in his *stochastic music*: "I proposed a world of sound-masses, vast groups of sound-events, clouds, and galaxies governed by new characteristics such as density, degree of order, and rate of change, which required definitions and realizations using probability theory"(Xenakis, 1992, p. 182). He applied his techniques to all the elements of music, from pitch, scale, and rhythm, to overall form. He also explored sound synthesis and timbre, from his orchestral work *Metastasis*(1964), down to one of the earliest approaches to granular synthesis in *Analogique B*(1958), where small short 'grains' of sound layered together form new, never before heard textures.

The computer naturally affords the definition formal procedures, and the speed of computation the realisation of complex algorithmically generated content, previously beyond reach. An ever expanding array of computer languages and systems would be created, and continue to be created, for application to music and composition(Loy, 1989; Fernández & Vico, 2013). Computer-assisted composition, also known as computer-aided algorithmic composition (CAAD) (Ariza, 2005a), would be an increasingly common practice as computers became readily available and affordable. In computer aided composition, formalism is explicitly a primary concern of the composer, enabling her to apply generative processes to multiple levels of musical abstraction, in a complex mix of programming, parameter selection and human evaluation.

The 'musicological research' side of computer music produced continual advances in tandem with advances in Artificial Intelligence. Software to imitate the styles of composers and musical genres grew in number and capability; harmonisation of chorales in the style of Bach, and modelling jazz improvisation particular academia favourites(Ebcioğlu, 1990; Hild, Feulner,

& Menzel, 1992; Allan & Williams, 2005; Phon-Amnuaisuk & Wiggins, 1999; Ebcioğlu, 1986; Biles, 1994; Toiviainen, 1995; Papadopoulos & Wiggins, 1998). This is nowhere more clear than in Cope's famous anthropomorphised *Emmy*. Built on a huge database of music of past composers, Emmy could generate new compositions in that style. Cope even had Emmy compose 'Beethoven's 10th Symphony', to no shortage of controversy(Cope, 2005).

Connectionist AI, whereby the symbolic instruction sets that programmers traditionally deal with are replaced by self-organising virtual structures, such as neural networks, have also been applied to both 'musicological research' and exploratory sides of computer music(Fernández & Vico, 2013). In this style of programming, the composer/programmer is obfuscated from the detailed workings of the software, and instead expose or 'train' a learning mechanism by exposing the networks to a desired behaviour. Further Artificial Life techniques, such as cellular automata and evolutionary processes would also be applied to music in both research and creative contexts. Artificial Life is re-visited in section 2.2.4.

Developments in sound synthesis with computers occurred in tandem with developments of compositional systems. Here, generative processes operate in shorter time frames, and rather than resulting in the creation of melodies, sound synthesis is concerned with the creation of timbres. Some of these techniques are based on digitally modelling acoustic phenomena, seeking to mimic real-world sounds and instruments(Strobl, Eckel, Rocchesso, & le Grazie, 2006; Jensen, 1999; Karjalainen & Smith, 1996), whereas others, such as Chowning's discovery of frequency modulation(Chowning, 1973) used the computer to create new sound worlds. Many such experiments in the artificial generations of timbres preceded the computer, and were built on a wide array of technological substrates. As Rhodes points out "the evolution of sound synthesis has always been interwoven with the engines of acoustic emission, be they mechanoacoustic, electromechanical, electrooptical, analog electronic, or digital"(Roads, 2004, p. 44). But with the computer came an increasingly sophisticated palette of signal processing tools. What was once the base unit of the traditional *lattice sonics* – the 'note' – itself would be subject to formalised processes, its morphological features manipulated by composer/programmers.

With the reduced cost of the technology, artificial synthesis of audio via digital means is a standard generative process used by practitioners of the vast and rich field of electronic music. With tools such as sequencers for higher level organisation, and components, physical or virtual, that can be wired and interconnected in infinite combinations, the electronic musician is embedded in layers of generative systems covering both micro and macro timescales. Today the maturity of this practice is well illustrated by the existence of 'live-coding' performances, whereby performer/composers would write software for both micro and macro scale algorithmic generations, live on stage in a performative context(Collins & McLean, 2014).

2.2.3 Visual Arts

In the previous section, it was suggested that for music, the definitions of conventions and style could be understood as 'generative' encodings. However this line of reasoning does not apply so easily to the history of visual arts. Here, style and conventions are rather related to conceptual considerations or methods of technical production, and are not so easily 'codifiable'. In the Western 20th century fine arts, when they are defined, they instead appear as wordy 'artists statements'. There is no 'lattice sonics' around which the visual arts have defined themselves. Further, images and paintings are capable to act a signifiers, representing 'things' in the world, whereas a 'note' does not inherently carry a meaning; instrumental music is not 'about' anything external to itself.

For these reasons, there is not the same level of codifiability in the core narratives of the history of western fine-arts, as one finds in music. However that does not mean that generative processes are not present. On the contrary, some will even contend that generative art is "as old as art itself"(Galanter, 2003). As is evidenced by the countless instances of iterative geometry and symmetry applied to textiles, drawing, sculpture, tiling, ornamentation and architecture; patterns have been present in human creative endeavours ubiquitously across cultures for millennia.

The *definition* of a pattern, its description, can be understood as an algorithm. Once a pattern is begun, like in a cannon, the output is 'generated'. The artist cedes an amount of control to an abstract system. As is the case for music, wether to describe an artefact as 'generative', is a subjective fuzzy (non-binary) description. Rather 'generative' is a description of an abstract process, a mode of creation, independent of media and technological mechanisms[2].

In generative design, a fundamental activity is the specification of the logic and systems that result in the patterns. Once the logic and rules are specified, these delineate the bounds of the design space within which the artists operate(Prats, Earl, Garner, & Jowers, 2006). The design space can take the form of tradition and style, the formalisms of which get passed across generations, evolve and mutate. Alternatively, artists or designers can, in a proactive and focused

[2] Any computer based generative method could, in principle, be carried out manually.

manner, be concerned with developing the systems and processes themselves, as is the case with computer art, discussed in section 2.2.4.

Patterned creation is described by Galanter as ‘highly-ordered’ generative art, the ubiquitous ancient procedures found across cultures; “The artistic use of tiling, in particular, is nothing less than the application of abstract systems to decorate specific surfaces.. the most notable examples in this regard are perhaps the masterworks found in the Islamic world”(Galanter, 2003).

Like other art forms, classical Islamic art does not exhibit a unified discourse; its styles were subject to the forces of historical changes in politics, theology and technology(Shalem, 2012). There is much uncertainty on the exact procedures by which the ancient Islamic tiling works were created, although evidence suggests that they are largely based on algorithmic and mathematical procedures(Özdural, 2000; El-Said, El-Bouri, & Critchlow, 1993). Jowers et al. have demonstrated that they can be described in terms of ‘shape grammars’(Stiny & Gips, 1972); a formalism for the generative specification of forms, supporting the view that such works are designed according to a “clearly-formulated method incorporating a coherent and adaptable system of composition allowing both variation and innovation”(Jowers, Prats, EISSA, & LEE, 2010).

The tiling works of the Islamic world are remarkable in their great phenomenal visual complexity. However they are constructed from simple elements and forms such as squares, triangles and stars. Then through procedural sequences of translation, reflection and recombination, complex forms and structures emerge. The emergence resulting from these inter-combinations of the elements, exhibit properties of symmetry, scalability and a visual complexity not readily perceived or anticipated when contemplating the work as a set of instructions. They have a latent internal organising logic whereby a ‘specificational simplicity’(Stiny & Gips, 1972) can nonetheless result in an emerged visual complexity.

There are numerous suggestions that this ‘emergence’ of the visual representation from the underlying structures and algorithms have theological motivations and considerations. Marks suggests that as the image is contemplated, its latent and internal logic is ‘unfolded’ by the viewer; the work embodying “a logic whereby the multiplicity of creation unfolds from the infinite unity of God... the visual image an expression of a divine ‘logic’ that may or may not be perceptible”(Marks, 2006). Jowers et al. support this view: “the patterns support exploration into philosophical aspects of Islamic faith, such as the relationship between unity and multiplicity”(Jowers et al., 2010). Gocer suggest that a unifying philosophical concern of Islamic art lies in a Platonic ideal: “A philosophical explanation of the Islamic attitude toward art may therefore be derivable from the Platonic axiom that god is the paradigm of beauty, which is the cosmological principle of order and harmony; and as all created things, art, too, must reflect the divine paradigm”(Gocer, 1999). These perspectives consider Islamic art as a confluence of art, mathematics, philosophy, and religious thought.

Other cultures have also shown evidence of generative processes within theological contexts. Hindu temple architecture for instance, exhibits distinctly self-similar iteration and fractal cascading forms. Kirti Trivedi suggests that the Hindu temple can be understood as ‘a model of a fractal Universe’: “According to Hindu philosophy the cosmos can be visualised to be contained in a microscopic capsule, with the help of the concept of subtle element called ‘tammatras’. The whole cosmic principle replicates itself again and again in ever smaller scales”(Trivedi, 1989).

Hindu temple constructions are based on ancient texts on architectural practice called *Vastu-shastras*, which describe procedural process for their design, layout and build. Trivedi points out that they show strong parallels to approaches of computer graphics “including discretisation, fractalization, extensive use of recursion and procedures involving self-similar iteration”(Trivedi, 1989).

One such instruction reads like this -

> “The layer of prahara (projection) will be above the chadya (eave of the roof). This is to be repeated again and again on the spire over the spire. A fraction of the prahara is to be constructed and again the spires are to be constructed. Each of the upper spires will be sprouted out with a measurement equal to half the size of the lower spire” - Ksirarnava, 7.113, quoted in (Trivedi, 1989).

Such a text transmits a recursive ‘program’ to be executed by masons. Initial decisions act as parameters, or seeds, that are fed into the rule sets, the resulting structure an emergent form; a kind of generative masonry. Sambit Datta further analysed such constructions and modelled the generative process in software. He was then able to algorithmically generate the self-similar fractal structures of 10th century Hindu temple architecture(Datta, 2010). He conjectures: “Considering the philosophical and mathematical concepts revealed by this method of reconstruction, were ancient Hindu temple builders grappling with a method for encoding a notion of infinity through their use of geometric sequences?”(Datta, 2010)

The *Vastu shastras* manuals were not limited to temples, and further included chapters on the constructions of residential and even military buildings, the layout of towns, gardens, roads, and water works, market-places(Acharya, 1946, p. xvii).

Needless to say, procedural elements can be found in cultural artefacts in other continents. Ron Eglash, in his book ‘African Fractals - Modern Computing and Indigenous Design’(Eglash, 1999), provides a useful categorisation of cultural manifestations of mathematical processes, placing them along a spectrum from the ‘unintentional’ to ‘fully conscious’:

> “At one extreme is the emergence of completely unconscious structures. Termite mounds, for example, are excellent fractals . . . we next find examples of decorative designs which, although consciously created, have no explicit knowledge attached to them. It is possible, for example, that an artist who does not know what the word ‘hexagon’ means could still draw one with great precision. This would be a conscious design, but the knowledge is strictly implicit . . . In the next step along our spectrum, people make these components explicit – they have names for the patterns they observe in shapes and numbers. Taking the intention spectrum one more step, we have rules for how these patterns can be combined”(Eglash, 1999, p. 5).

This more intentional end of the spectrum would include the examples of Islamic tilings and Hindu generative temple structures; the knowledge is externalised in a body of cultural practice passed along through tradition. These also correlate well with the trajectory of the history of music discussed in the previous section; the conventions of styles in music becoming more codified and explicit, even more so as serialism, and eventually computer based algorithmic composition practices took hold.

Eglash provides numerous instances of increasingly explicit generative practices in African cultures, with a strong prevalence for fractal structures as a ‘design theme’ that has spread across the continent: “not only in architecture, but in traditional hairstyling, textiles, and sculpture, in painting, carving, and metalwork, in religion, games, and practical craft, in quantitative techniques and symbolic systems, Africans have used the patterns and abstract concepts of fractal geometry”(Eglash, 1999, p. 7).

Returning to European and Western cultural practices, formal processes can often be found in folk art, ranging from fractal Celtic knots(Kaplan & Cohen, 2003), intricate geometries in Ukrainian *Pyansky* Easter eggs(Danko-McGhee, 1999), to the self similar abstractions of Ammish quilts(Smucker, Shaw, & Cunningham, 2009). Perhaps the most developed practices of patterned design lie in the ‘fibre arts’ (knitting, crochet etc) whereby the substrate is fundamentally algorithmic and mathematical.

Outside the arts, the industrial revolution brought about physical generative mechanisms in the form of machinery as a replacement for hand production; manufacturing can be considered a form of generative design. One key breakthrough was the invention of the Jacquard Loom in 1805. As one of the first ‘programmable’ machines, it augmented textile machines with parameterisation. It enabled the sequence of operations to be defined by replaceable punched cards, and in introducing the notion of a stored ‘program’, it is considered a key development in the history of computing(Fernaeus, Jonsson, & Tholander, 2012).

Within the context of Western fine-art however, generative practices would not appear until the 20th century, as the emphasis on mimetic pictorial representations faded, and the use of abstraction would become increasingly prevalent. Spiritual concerns motivated some of the earliest abstract works, as seen in for instance in the work of Hilma af Klimt(Fant, 1986). Similarly Emma Kunz and Agnes Martin created highly patterned works of considerable geometric complexity within a context of a spiritual belief system. Catherine De Zegher describes the works of these three artists as analogous: “Engaging in the repetition of certain distinct geometrical shapes (square,triangle,circle etc..) marking the experiential, the spiritual and the infinite through recurring structures of linear patterns, the three artists believed that abstraction might be a means for achieving higher cognitive levels inextricably linked to the forces and processes of life”(af Klint, Kunz, & Martin, 2005, p. 24). Their spiritual motivations have similarities to those of classical Islamic art and Hindu temple architecture discussed earlier.

Agnes Martin would feature in the aptly named exhibition *Systemic Painting* at the Solomon R. Guggenheim Museum in 1966, alongside the works of artists such as Carl Andre, Peter Gourfin, and Will Insley(Alloway, 1966). Many of these works contained ordered geometric pattens created through procedural and numerical methods, and helped define the ‘minimalism’ art movement. By this time, patterns and geometric form would become increasingly common in the art galleries. Other notables including Donald Judd, Robert Smithson and Sol Lewitt, all of which used geometric and numerical systems as generative elements(J. Meyer, 2004). Further the Op-Art movement, as exemplified by Bridget Riley, also carried an emphasis on highly patterned visual structures, as did the works of M.C. Esher, in his geometrical “regular divisions of the plane”(Locher, Bool, & Ernst, 1992).

The level at which these artist would externalise their processes, and how ‘implicit’ or ‘explicit’ the generative procedures would vary for each style, artist and individual work. And as in music, wether to call a work an instance of ‘generative art’ is a fuzzy consideration. For some theorists such as Galanter(Galanter, 2003), and Bodens and Edmunds(Boden & Edmonds, 2009), the key consideration is that some element of artistic control is relinquished to an external system of some sort. ‘Generative art’ was not known as an explicit art movement, and rather just represents a way of working, as Galanter points out - “Generative art is ideologically neutral. It is simply a way of creating art and any content considerations are up to the given artist”(Galanter, 2003).

Though chance elements had already been at work decades earlier (Mark Tobey and Jackson Pollock famous examples), previously chance occurred in the immediate rendering of the pieces (e.g. drips of paint), rather than as part of an externalised autonomous process. Sol Lewitt famously made explicit generative processes in his 1967 ‘Paragraphs on Conceptual Art’, where he stated the art should be created by formulaic rules: “The idea becomes a machine that makes the art” where “all of the planning and decisions are made beforehand and the execution is a perfunctory affair”(LeWitt, 1967). His ‘Instruction-based Art’, included chance elements as part of the generative principles, leaving decisions to be made by the interpreters of the instructions. This is a similar sentiment to some of the aleotoric works of avant-guard composers of the 50s and 60s, such as Cage, Stockhausen, Earl Brown and Morton Feldman, whereby elements of the resulting music were similarly left at the performer’s discretion.

2.2.4 Computer art

In the 20th century, there are parallels in the historical developments of generative art and music. As artists would bring in chance to art-making, so would composers. And as computers came into the picture, the most common early generative mechanisms explored were stochastic processes, in both music and computer art. The first exhibition of computer art was called *Generative Computergraphik*(Stuttgart,1965) and displayed the work of Georg Neese. It was closely followed by exhibits showing Frieder Nake and Micheal Noll(Klütsch, 2007). Those three artists were the key figures of the dawn of computer art, and shared similar approaches. As Frieder Nake states: “Almost all early experimenters in visual computer art made use of random numbers.. If we were talking about natural art (the opposite of artificial art), we would interpret such an event as style: the common manner by a group of artists to draw or paint”(Nake, 2005).

These early works have a very similar aesthetic, often consisting of monochromatic minimal line-based compositions, and as Nake points out, employed very similar methods. He describes the approach as a distinction between *macro-aesthetics*, representing the high -level form and overall probability distributions, and the *micro-aesthetics*, the local details of the work: -

> "When we look at a work of computer art, and analyze its overt structure, we may in many cases fairly well describe the global geometry, which here I call macro-aesthetics. We are at odds, however, with the details that, as we know, are decided by random numbers. These details on the local scale constitute the micro-aesthetics. They remain hidden to us. The same is true for the artist's intuition"(Nake, 2005).

These early pioneers were influenced by *information aesthetics*, a scientific approach to the aesthetic experience, drawing on Shannon and Weaver's information theory, as well as Norman Weiner's cybernetic theory and other formal theories, such as Birkoff's measures of aesthetic pleasure[3]. Information aesthetics was championed by Max Bense, who curated the first computer art exhibition and mentored George Nees(Klütsch, 2007; Nake, 2005). In fact, the very first computer generated drawings were created to explore and test Bense's aesthetics(Klütsch, 2012).

Nake reflected on 'information aesthetics' and its influence to early computer art -

> "Information aesthetics attracted considerable attention because of its scientific attitude. All romanticism was banned from it. No subjectivism was left. An aesthetics of the object was intended that should come close to the way physicists were studying their subject"(Nake, 2005).

There are here echoes of Xenakis' enthusiasm of the application of scientific methods and thought to music.

In having brought the use of the computer to art-making, there was as in music, a new extreme emphasis on the formal generative properties of art-making. The generative process itself would become the aesthetic concern, representing a *class* of works, rather than concrete individual objects.

Although there was considerable effort in early generative music to model existing genres, there was not as much emphasis on mimesis in early years of visual computer art[4]. One notable

[3] $M = O/C$, where the aesthetic measure M is defined as the ratio of order O over complexity C. Birkhoff's measures is re-visited in section 2.6.3

[4] Although decades later, some researchers would look to 'evolve' images in the styles of certain artists.

exception was carried out by Michael Noll. His study of Mondrian's *Composition with lines*, an abstract minimal work consisting of distributions of short vertical and horizontal lines, pinned computer generated emulations against the original in a preference test: "Only 28% of the [subjects] were able to correctly identify the computer-generated picture, while 59% of the [subjects] preferred the computer-generated picture"(Reichardt, 1970). Nake's oft cited 'Hommage á Paul Klee, 13/9/65 Nr.2'(1965) has been commented to be "a literal analysis of an oil painting by Paul Klee.. to explore different visual effects, based on Klee's 'repertoire' of imagery"(Beddard, 2009). Although having stated that he was inspired by Klee, Nake denied that there was any real attempt to emulate these works - "I myself didn't do enough against this misinterpretation"(Nake, 2005).

The evolving sophistication of generative techniques would coincide with developments in the sciences. In addition to Shannon's information theory, which, as already discussed, was hugely influential for both generative artists and composers, cybernetics is another area of research that has been instrumental in the development of generative music and art.

In 1948, Norbert Weiner defined cybernetics as the scientific study of "control and communication in the animal and the machine"(Wiener, 1961). It studies the ways in which mechanical, biological and social systems organise themselves, regulate themselves, reproduce themselves, evolve and learn(Pask, 1961). It views both living and non-living entities as composed of layered systems with inputs and outputs that reacted to, and act upon, each other and the environment in complex feedback loops; 'self-organisation' one of the field's core concerns.

Cybernetics comes from the greek *kybernetes* - meaning steersman, governor or pilot(Wiener, 1950). Implied in the etymology of the word is the concept of the *cybernetic loop*. Here, a system and an environment are differentiated by a boundary. The system acts, senses results, compares to its goal. In other words the steersman seeks to guide the ship along a path, sees which way to adjust the rudder, acts by steering left or right, gets feedback from the environment, and so on. From a cybernetic perspective, generative design involves a *navigation* or search through parameter space; the artist evaluates an output, adjusts parameters, evaluates output, adjusts parameters and so on. The feedback loop a defining characteristic of the interaction between the artist and the system.

For instance Schnier and Gero used evolutionary algorithms to emulate paintings by Mondrian and Frank Loyd Wright window designs(Schnier & Gero, 1998), and Eiben would evolve Mondrian and Escher style works(Eiben, 2008)

Cybernetics, and its related fields, brought new understandings of the complex, dynamic systems that underly natural processes. These understandings would lead to new models and theories, that in turn would be appropriate by artists and composers as generative techniques and systems. These range from neural networks, and cellular automata, to reaction-diffusion systems, swarm and flocking simulations, and chaotic systems.

The seminal exhibition *Cybernetic Serendipity* in 1968, described by its curator Jasia Reichardt as "dealing broadly with how man can use the computer and new technology to extend his creativity and inventiveness"(Reichardt, 1970), represented a fusion of all of the arts with technology. The accompanying booklet, *Cybernetic Serendipity – the computer and the arts*, documents a broad range of application of generative processes to all arts; visual, music, text, dance, architecture, sculpture, and featured a 'who's who' of computing, composition and art pioneers.

In its pages, Norman Weiner and Stafford Beer explicate core concepts of cybernetics. One reads histories of computer development and theory, texts on computer music by Joseph Schillinger, explanations of stochastic musical processes by Stockhausen, and Hillier outlines his 'musicological research' behind the *Illiac Suite*, alongside numerous descriptions of other early computer music systems. One also finds John Cage's score for the non-deterministic tape work, *Fontana Mix*, together with some zen musings on computer art and the nature of 'the audience', as well as outlines of a 'computer-programmed choreography' for human dancers. Cybernetician Gordon Pask describes the 'aesthetically potent' environment (now commonly known as the interactive art installation) in his presentation of his *Colloquy of Mobiles* - where groups of mechanical mobiles, through light flashes and sound, affect each other in an intricate generative choreography, and that could further be influenced by the audience. Inhaowitz, famous for his *Senster* sculpture, showed an early work,"an electro-hydraulically operated, environment sensitive mobile". One learns about various drawing and painting machines and programmes, light sculptures, generative texts and poetry, computer generated video, and even computer generated architecture. Also included were the works of the three afore mentioned pioneers of computer graphic art, Nake, Hasse and Noll (where Noll presented his previously mentioned study of Mondrian).

As the computational practices took hold and spread, an ecosystem of art genres sprung forth, coupled with a colourful taxonomy, as Boden and Edmunds summarise:

> "The names preferred by the artists involved include: generative art, computer art,

> digital art, computational art, process-based art, electronic art, software art, technological art, and telematics. All of those terms are commonly used to denote the entire field–and (although distinctions are sometimes drawn) they are often treated as synonyms. In addition, there are names for subfields: interactive art, evolutionary art, video art, media (and new-media and multimedia) art, holographic art, laser art, virtual art, cyborg art, robotic art, telerobotics, net art ... and more. Again, the extension of these labels is not always clear"(Boden & Edmonds, 2009).

The *Cybernetic Serendipity* exhibition represented a fusion of techniques in generative art and music, and applications of these techniques to other art forms; poetry, dance and architecture. As exemplified by the audio/visual sound installations, hybrid media would become increasingly prevalent in technology-mediated art (the term 'new media art' sometimes used as an umbrella). Furthermore, *interactivity* would become a new paradigm, with an increasing emphasis on making works that would sense and react to contemplators. Artists such as Roy Ascott and Nam June Paik would create works and interventions mediated by the principles of feedback between human and technological systems. Technology was now more than just the means by which a work was created: technology, and society's relationships to it, would be the subject of aesthetic consideration. As Whitlaw points out: "New media art self-consciously reworks technology into culture, and rereads technology as culture"(Whitelaw, 2004).

Space limits considerations of continuing a detailed historical survey of the applications of all the developments of 'new media art', however two key sub-areas will be discussed: evolutionary art and music (EMA), and a related branch of research that explores the generative potential of feedback loops - dubbed here 'ecosystemic' works.

Evolutionary Art and Music

Since the early days of computing research, a number of techniques inspired by generative process of nature were developed. These include Lindenmayer-systems (or L-systems), a mathematical theory of plant topology developed by Aristid Lindenmayer in 1968(Prusinkiewicz & Lindenmayer, 1996). His L-systems are a computational model of the fractal branching structures of plants, which could be used to synthesise plant forms.

Another early technique developed in the early days of computation are neural networks. First explored in the 1940s, they seek to emulate neural processes in the brain, and have been applied to machine learning, classification, and control tasks ever since(Lippmann, 1987).

Reaction-diffusion systems, initially explored by Alan Turing(Turing, 1952), were developed as a prototype model for pattern formation and the 'morphogenisis' of biological forms. Similarly, cellular automata, conceived by Stanislaw Ulam and John von Neumann in the 1940s and 50s in an attempt to understand the properties of self-replicating systems, generate global high-level behaviours based on the interactions of simple units(Von Neumann, Burks, et al., 1966).

Darwinian evolution inspired a whole subfield of artificial intelligence, called 'evolutionary computation'. Computational evolutionary processes were first explored in the 50s and 60s, and have yielded a whole class of programming techniques readily applied to optimisation problems in engineering(Back, Fogel, & Michalewicz, 1997). Key to evolutionary computation processes, is how the fitness of an artefact is determined. A computationally defined *fitness function* can act as an objective measure of how 'fit' an artefact is, and hence how well it will fare in the evolution.

More recently agent-based systems explored 'swarm intelligence', inspired by the collective self-organising behaviours of birds and insects, in the form of flocking algorithms and ant-colony optimisation systems(Engelbrecht, 2006).

A core theme that cuts across these techniques is the *emergence* of self-organisation, as low level elements interact to yield macro-level behaviour and forms. For instance, in cellular automata, such as in Conway's *Game of Life*(Conway, 1970), each cell follows very simple rules, but across the 'ecosystem' complex behaviours manifest themselves. Similarly in flocking algorithms, such as Reynolds *Boids*(Reynolds, 1987), each 'boid' is no more complex than a thermostat, following very simple rules, but when the environment consists of numbers 'boids', complex behaviour emerges in the form of swarms.

Although these techniques initially were discovered and studied as part of scientific enquiry, artists and composers would apply these techniques, and combinations and variations of these, to the creation of visual, sonic and multi-media works.

For instance cellular automata have been applied to music both at the macro (note) level (e.g. Xenakis *Horos* (Solomos, 2006)) and the micro (timbre) scales (e.g. Miranda's *Chaosynth* (Miranda, 1995)), as well as to visual computer art (e.g. Paul Brown's *Sandlines* (Brown, 2001)). Other artists made use of agent-based systems, such as flocking and swarm simulations, both in audio synthesis and music(T. Davis & Karamanlis, 2007; Blackwell, 2008), as well as visual art(Moura & Ramos, 2002). L-systems have been appropriated by artists such as Jon McCormack to the creation of forests of virtual plant forms(McCormack, 2003). Others have applied L-systems to the generation of musical content(Worth & Stepney, 2005; Manousakis, 2006).

Dawkins *Biomorphs*(Dawkins, 1986) is the first well-known instance of the application of evolutionary computation to computer graphics. Here human users would apply a selective pressure by choosing which virtual entities would survive in an evolutionary process, based on their appearance when drawn on screen. Dawkins' concerns were scientific – to demonstrate concepts of Darwinian evolution by showing how small incremental changes could lead to complex forms and behaviours – but it would not be long before they inspired others to use these techniques with aesthetic motivations.

Sims(Sims, 1991) and Todd and Latham(S. Todd & Latham, 1994) applied the techniques of evolutionary computation to the creation of images and virtual creatures for aesthetic ends. In evolutionary algorithms the biological concepts of *genotype* (the genetic information), *phenotype* (the individual itself), *expression* (the process whereby a genotype is used to construct a phenotype), *fitness* (the entity's likely hood of survival due to its attributes), *selection* (how the fitness of the phenotype is determined), *reproduction* (how new entities are created from existing surviving phenotypes) and *variation* (how mutation is applied to entities to ensure variety in the gene pool), are all modelled in software(Sims, 1991). However selecting parameters for aesthetic ends is not straightforward: there is no 'fitness function' for beauty[5]. Instead selective pressure would be driven by human judgments; the fitness of the artefacts (phenotype) determined by their aesthetic properties. Typically in the works of Sims and Latham, the state-space, the range of possible images, would be vast (or unbounded), resulting in open-ended and difficult to predict images.

Since then, variations of the above techniques, sometimes in combination, have been readily applied to the creation of visual and sculptural content(Whitelaw, 2004; J. J. Romero & Machado, 2007; M. Lewis, 2008; C. Jacob & Hushlak, 2008), as well as to music(Dahlstedt, 2004; Miranda & Biles, 2007) by artists, composers and researchers. A related approach explores the application of ecosystem simulations to create artistic artefacts(Dorin, 2008).

There are difficulties with having humans perform aesthetic selections in evolutionary computation; practically only a small portion of the possible state-space can be explored because humans are too slow and cannot evaluate many outputs at once. Additionally cross-breeding ap-

[5]Although there have been attempts at formalising aesthetic properties of objects (some of these are discussed in 2.6.3), a general computable fitness functions relating to human aesthetics remains an open and intractable problem.

proach to the selection of parameter-sets can result in large jumps in state space and difficult to anticipate outputs(McCormack, 2005).

With the increased ubiquity of the internet, crowd-sourcing – also known as ‘human-computation’ – whereby numerous volunteers carry out tasks towards a common goal, has been increasingly explored by researchers(Quinn & Bederson, 2011). Crowd-sourcing has been applied to numerous goals, from image search and recognition tasks to the crowd-sourcing of aesthetic designs(L. Yu & Sakamoto, 2011; Nickerson, Sakamoto, & Yu, 2011). It is today a common approach to eliciting feedback for interactive evolutionary computation (e.g. (Kosorukoff, 2001; MacCallum, Mauch, Burt, & Leroi, 2012; Draves, 2008)).

In crowd-sourcing, the judgments of the individuals are merged in a *wisdom of crowds*(Surowiecki, 2005). Thus the interactive evolutions necessarily represent a kind of ‘average’ of the aesthetic preferences, and if aesthetic judgments are both subjective and objective (see section 2.6), only the objective elements will come through in these processes.

Two of the practical studies of this thesis apply crowd-sourcing to the collection of aesthetic judgments. The *Melody Triangle* as a mobile phone application collects musical preferences from its users, and *EvoColour* applies evolutionary computation to the evolution of colour palettes, the subjects of chapters 3 and 5 respectively.

Evolutionary computation is extensively used in architecture and design. It is a well established technique for evolving designs to match structural engineering requirements, such as efficient use of materials, as well as for searching for novel aesthetic forms(Frazer, 1995; Hensel, Menges, & Weinstock, 2004; Hemberg et al., 2008).

Miranda similarly identifies engineering and creative goals to the use of evolutionary computation (EC) in music: “Whereas the engineering approach applies EC to solve specific engineering problems in Music Technology, the creative approach employs EC to generate musical compositions”(Miranda, 2004). Examples of the ‘engineering’ problems could be finding optimum parameters for a synthesiser(Horner, Beauchamp, & Packard, 1993; Mitchell & Pipe, 2005; Garcia, 2001), where as the ‘creative’ approach considers the new musical possibilities afforded by these processes(e.g. (Eigenfeldt & Pasquier, 2013; Hoover, Rosario, & Stanley, 2008; Dahlstedt, 2004; Waschka II, 2007)).

In addition to the engineering and creative motivations, Miranda identifies the ‘musicological approach’, which “seeks answers to musicological problems by means of models and simulations"(Miranda, 2004). Such studies often make use of virtual agents to study how musical styles and languages evolve in communities of agents(P. M. Todd & Werner, 1999; Manaris et al., 2007; Miranda, 2002, 2004; Kirke & Miranda, 2015). Similar approaches have been also applied to communities of virtual agents evolving visual artefacts, as an exploration of how artistic styles evolve and are reinforced, often using virtual 'art critics'(Greenfield, 2008; Galanter, 2012; Machado, Romero, Santos, Cardoso, & Pazos, 2007; Machado, Romero, Santos, Cardoso, & Manaris, 2004; J. Romero, Machado, Santos, & Cardoso, 2003).

Interactive evolution has also been studied by MacCallum et al. to trace the role of consumer preferences on the formation of musical structures and style(MacCallum et al., 2012).

As these various, nature-inspired, computational techniques matured, they would be grouped together into a field called *artificial life*, or *a-life*. A-life, is a branch of both scientific and artistic enquiry, concerned with the simulation (or synthesis) of living things. It looks upon life as an emergent consequence of the interactions of low level elements in a complex dynamic system. Unlike 'good old fashioned AI', which attempted to synthesise life-like properties from the 'top-down', a-life instead considers a 'bottom-up' approach. As Christopher Langton, the founding figure of the field put it: "life emerges out of the organized interactions of a great number of nonliving molecules, with no global controller responsible for the behaviour of every part"(Langton, 1989, p. 3). Life is understood as an abstract dynamical process, and as such need not be exclusive to a biological substrate. A-life proposes that life, in principle, could be manifest in an artificial medium, such as software: "By extending the empirical foundation upon which biology is based beyond the carbon-chain life that has evolved on Earth, Artificial Life can contribute to theoretical biology by locating life-as-we-know-it within the larger picture of life-as-it-could-be"(Langton, 1989, p. 1).

Computational techniques that sought to model natural process were developed long before Langton's delineation of the field. However with the definition of a-life, the goals of engendering an emergence of life-like behaviour and forms were made explicit. The use of generative processes that mimic life represents a pinnacle of generative art practice. Here, the artist attempts explicitly, to make works that make themselves, that create themselves. As Whitelaw puts it, it is creating creation - *Metacreation*. He asks a pertinent question: "Is this an abdication of creative will or its ultimate fulfilment?"(Whitelaw, 2004, p. 3)

'Ecosytemic' works

Before completing this survey of generative practice in the arts, one last sub story will be considered here. Although not considered a genre of art per-se, there has been an ongoing thread of activity to explore the aesthetic properties of feedback systems, often influenced by the core principles of cybernetics.

A number of artists and composers have to engender emergent behaviour by setting up complex, non-linear feedback loops. These works consist of one or more entities, reacting to and acting in an environment, perhaps including humans in the loop. The cybernetician Gordon Pask created a number of works along these lines. This includes his electrochemical *Ear*, a device with a kind of structural autonomy created in the 1950s(Cariani, 1993), and his previously mentioned 'Colloquy of Mobiles'. Inhawotiz made a number of 'cybernetic sculptures' that would react to sound or light, his famous 'Senster' an other example. Alvin Lucier's *Music On A Long Thin Wire* is an autonomous sound sculpture, its sounds continually changing in an unpredictable, organic way due to the complex non-linear feedback of acoustics and resonances(Harder, 2012). Along similar, lines Agostino Di Scipio's *Audible Ecosystemic Interface*(Di Scipio, 2003) uses the response of the surrounding acoustic environment to drive interlinked networks of control and audio signals.

Others have explored this paradigm as an augmentation of musical performance. In *Hornpipe*(1967), Gordon Mumma made use of a 'cybernetic console' that continually adjusted its signal processing in response to the sonic events in the performance space(Hiller & Isaacson, 1959, p. 101). More recently Simon Waters' *VPFI* (Virtual/Physical Feedback Instrument) Flute extends this paradigm, he develops the idea of 'Performance Ecosystems'(Waters, 2007) as an aesthetic and analytical orientation. George Lewis' *Voyager*(G. E. Lewis, 2000) system acts as a fellow performer for improvisation that responds to the actions of musicians and attempts to create a symbiotic relationship between system and musician.

Such an ecological approach attempts to remove any externally applied will or pre-defined final state; the behaviour of the system at any moment not reducible to an analysis of its parts in isolation. The form and behaviour *emerge* as the component affect each other through, and together with, the environment. There is here a unique model of interactivity; the standard view of 'user acts - computer reacts' where the user is the only source of creative, dynamic behaviour, is not applicable. Instead something akin to the kind of interactions living organisms normally have with their environment is manifested. Di Scipio says that such work is "a shift from creating wanted sounds via interactive means, towards creating wanted interactions having audible traces"(Di Scipio, 2003), or as Owen Green points out in his analysis of Di Scipio's *Audible Ecosystemic Interface* "interactions become the very subject of composition"(Green, 2008).

There are similar metaphors at work in both this ecological paradigm and in the evolutionary processes described earlier. However one important distinction lies in the system's interface to its users or environment. In ecosystemic works, there is no clearly differentiated, distinct steps of selection, evaluation, selection and so on. Further this differs considerably from computer mediated composition and art process, consisting of the distinct steps of defining the formalisms, populating it with parameters and evaluating the output. Rather outputs are the result of smooth, *continuous* feedback processes. Designer's in such situation engage in a process of 'tuning' the different elements and connections of the system.

This thesis will present an experimental eye-tracking interface that follows this 'ecosystemic' paradigm of continuous feedback, as a way of exploring the parameter-space of a generative systems, in chapter 4.

2.3 What is a Generative Process?

This thesis is largely concerned with generative processes and systems. As has been illustrated in the previous sections, generative processes are a subject of interest for artists, composers, designers and scientists. However due to the breadth and cross-disciplinary nature of this interest, the terminology and concepts to discuss and describe such processes vary, in particular when in artistic contexts. This section will draw on information theory and define some key terms that will be used throughout the thesis.

To begin the discussion on the nature of generative processes, a concrete example is provided. Consider the following pseudo-code for the generative process, G :

1. Draw a square, with edges of unit length.
2. Shift the canvas one unit length to the left.
3. Rotate the canvas 45 degrees to the left.
4. Go to 1.

Executing this algorithm results in this:

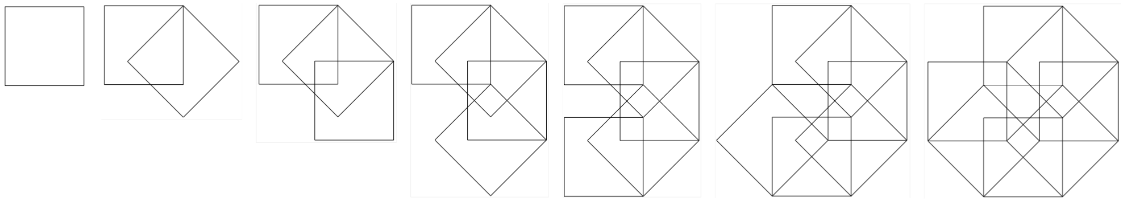


Figure 2.1: First 7 iterations of G.

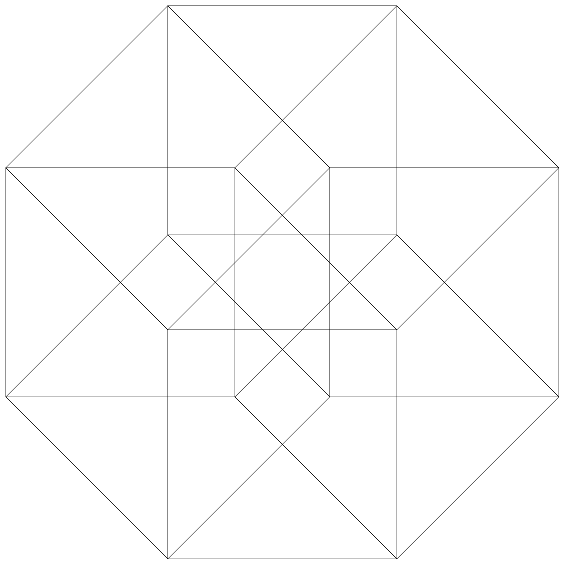


Figure 2.2: Output of G.

As can be seen in fig. 2.2, from the 9th iteration onwards, the squares are then drawn exactly on top of each other.

In the final output, numerous forms can be perceived; multiple stars, an octagon, even four Necker cubes are present. What is ‘seen’ at any one moment is the viewer’s interpretation. Further, presenting the result in a different rotation, primes the viewer to more readily see one interpretation over another (in the rotated form in fig. 2.3 the star is most readily seen). The ambiguous Necker cubes themselves engender shifts in perspectives; sometimes a particular face of the cube is at the foreground, at others it as the background. The human visual system is picking one interpretation or the other, despite the drawing being essentially just a set 32 lines of equal magnitude. These higher level forms are subjective interpretations carried out by an observer, not intrinsic attributes to the description of the generative process G. Such an example suggest a *pro-active* view of perception.

This resonates well with the famous Gestalt psychology credo ‘the whole is other than the sum of the parts’ which argues that perceptions cannot be dissected into basic elements. Cogntive

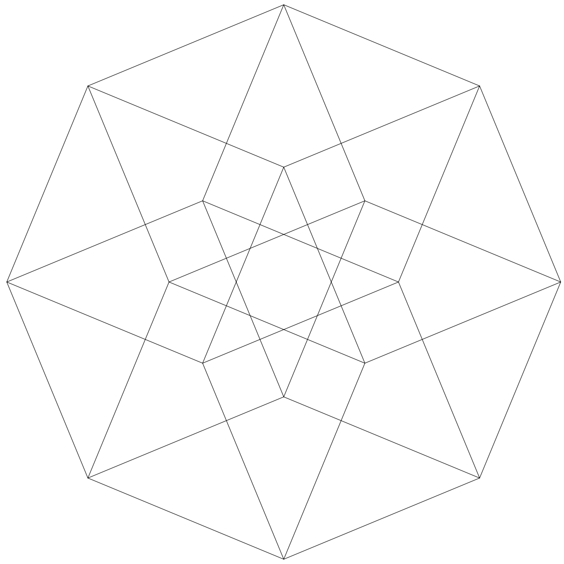


Figure 2.3: Output of G, rotated 45 degrees.

science suggest that both *bottom-up* processing, whereby knowledge is acquired from the senses, and *top-down* processing, whereby knowledge can influences sensations, are a part of perception. That the viewer can 'will' certain interpretations over another illustrates how this top-down processing is at play.

As is illustrated, the simple generative function G, returns an artefact that is perceived as more complex than the instructions *per-se* seem to embody. There is nothing in the instructions that suggests stars for instance. An experienced programmer or generative artist may have some intuitions on what the resultant form may result like just from reading G. However it takes some effort on the viewer's part to perceive the resultant form as the 8 squares that composed it, while not seeing the stars that resulted as a consequence of the intersections. The appearance of such higher level phenomena is often called *emergence*, a term that is prevalent in diverse scientific fields such as cybernetics and systems science, as well as in the literature of computer art. As Whitelaw states, this is particularly the case for a-life: "a-life art is defined by its constant evocation of an emergent result"(Whitelaw, 2004, p. 215).

Emergence can be understood as a shift in the level of description. From the descriptions at the level of squares, in can be said that there is an 'emergence' to the level of description of stars and cubes. Further, it just so happens that this pattern in fig. 2.2 is a 2d projection of a *tesseract*, also known as a hypercube, or a 4-dimensional cube. This is a level of description that however is not readily perceivable; this object is not easily understood as the 'shadow' of a 4 dimensional object.

Similarly, one can go in the opposite direction, and describe the artefact in lower-level terms than those used by the generative function G. This same form could be described as a set of 32 lines of equal magnitudes, or further still, if drawn on a screen, as a set of binary pixels on a screen. As such, the generative function G can have alternate formulations. For example G can be re-formulated with a lower-level set of instructions - G_2:

1. move forwards unit distance
2. turn 90 degrees to the right
3. move forwards unit distance
4. turn 90 degrees to the right
5. move forwards unit distance
6. turn 90 degrees to the right
7. move forwards unit distance
8. move backwards unit distance
9. turn 45 degrees to the right
10. go to 1.

A generative process can be understood as a set of instructions to be carried out in some medium, but it is abstract and exists independently of the medium in which the rendering occurs. The same principles can be applied to images, video, sound, sculpture, dance, kinetic sculpture; the instructions above could be carried out as some form of ritualistic dance, for instance.

Instructions that define generative processes are carried out, or rendered, by an interpreter. In computer based forms, this is the technological substrate; the software libraries residing on top of the hardware, with the instructions articulated in the form of a computer program. However this interpreter can just as readily be a human; for the above example the result can be carried out either manually or in software.

In generative art and music, it is humans who consumes the resulting artefact, and who perceive the emergent forms in the output. The computer does not ‘know’ about stars. However if it executed G, it would need to have some routine for drawing a square, or if G_2, some routine for drawing a line (as well as routines for rotation and movement). Those routines in a computer would themselves need to be programmed as a set of instructions, themselves built on top of lower level instructions down to the 1s and 0s that are interpreted as machine instructions.

For the example under consideration, this list shows different levels at which the generated output can be interpreted when viewed -

1. As points in a grid of pixels/ink on paper (the medium)
2. As 32 equidistant lines
3. As 8 squares
4. As 4 Necker cubes (alternatively various stars, shapes in an octagon)
5. As 1 2d projection of the tesseract/hypercube.

As discussed, in this case the level of the instructions for the generative process (3 for G, and 2 for G_2) is lower than the level at which the output is perceived (level 4). Note that the number of entities that compose the form decreases as one goes up the levels of description - $32 > 8 > 4 > 1$.

That there is a higher level of interpretation (level 4), that is not readily perceived is a point worth considering. Not all generative forms, that although may be described clearly in mathematical terms, can readily be comprehended in those terms. This may relate to levels of abstraction, as in the case of the tesseract, but more commonly with regards to computer generated artworks, relate to limitations of scale.

As a simple example, suppose a composer defines an algorithm to generate a melody that is cyclical and repeats perfectly. If the period of this repetition is too large, lasting weeks as an extreme example, then a listener would not have the requisite cognitive capabilities to perceiver this high level pattern. Similarly if a melody was composed to be palindromic, or to follow some strict logical procedures, such as those defined in serialism, it does not follow that the listener will hear the melody as palindromic, or be able to appreciate the logic of the serialist process. In other words, just because some algorithmic process has generated an output with some high level structures or is the consequence of intricate internal logic, it does not mean a contemplator will be able to grasp (or enjoy) the consequences of this logic.

To better explicate the workings of generative systems, the tools and language of information theory, drawing also on concepts from cybernetics, will be used throughout this thesis. This approach will first be used to describe generative process like the one above, and then will be generalised to include parameterised generative processes. These will be built upon further throughout the thesis to discuss self-organisation and emergence, and will be applied to the three practical studies; the *Melody Triangle*, *Keyebernates* and *EvoColour*.

2.3.1 Generative Processes as Constraint

As discussed in the review of generative music and art, generative processes can be viewed as restrictions on possibilities, as a set of *constraints*. Generative systems define a structured space of possibilities, but also restrict the range of perceived features that the output artefact can possess. Once a generative processes or design rules or conventions are set, the generated outputs will be of a 'class' defined by the process. The conventions of serialism, for instance, ensure that the music sounds 'serialist'. The rules and conventions of a fugue form ensures that the music sounds like a fugue.

A medium - such as a canvas, a score, or a computer display - can be understood as a *system* embodying a large space of possibilities. Systems have certain variable features, that can take on different values. All of the values for the different variables together determines the *state* - s - of the system.

The set of all possible states of a system is called the state space: $S = \{s_1, s_2, s_3, ...\}$. Where each s is defined by the same number of variables features - $v_1, v_2, v_3, \ldots$. At any one point in time t, if the system is not completely static and resistant to change, it may undergo a *transition* from one state s_t to the next s_{t+1}.

The 'output', the resultant artefact, is a *rendering* of one of the states of the system. For instance, that may be the process of turning the in-memory representation of the pixels into light on a display, or the realisation of an algorithmically generated score into sound, either by human players or digitally. In some cases there is no specific 'end state', and instead a sequence of states becomes what is rendered, such as frames in a generated video. Further this could happen in real-time, where the rendering occurs as the states are undergoing transitions.

Consider the case of a computer display of 100x100 pixels, each of which can be either black or white. Here each state consists 10000 binary values (representing wether a particular pixel is black or white). The total number of possible states - the state space - is astronomical (2^{10000}). This represent the sum total of all possible images composed only of black or white pixels at a modest resolution, this is for all purposes a range of possible outcomes too large to fully comprehend.

Any one possible outcome, or state will have a likelihood of occurring, this is the *probability* $P(s)$. The function P attaches a probability to each possible state, and as such determines a probability distribution over the state space S. When nothing is known about the system, then all states are equally likely - all states have the same probability P(s), yielding a homogenous probability distribution. Formally $P(s) = P(s')\forall s, s' \in S$. However, if everything is know about a system, then it is known with certainty that the system is in one particular state s. The probability in being in s is 1, but 0 for all other states. Formally $P(s) = 1, P(s') = 0 \forall s \neq s' \in S$.

Given the probabilities of our states, it is possible to quantify the degree of uncertainty using Shannon's *Theory of Information*. This uncertainty, H is known as the *entropy* of the system -

$$H(S) = -\sum_{s \in S} P(s) \log_2 P(s) \tag{2.1}$$

H is maximal when all states have the same probability. Conversely $H = 0$ when one state has probability 1, and all other states have probability 0. H can be understood as a measure of an observer's ignorance about the system's state.

Knowledge of the probability of being in one state P(s), is a *subjective* value, and is representative of how much information is known about the system. To someone who knows nothing about a system, for instance a visitor to an art gallery to view the canvas S with no prior knowledge, $H(S)$ is maximal before they view it[6].

However, the process of applying a deterministic process such as G, *given* a starting state s_0 where all pixels are white, there is only one possible result, s_g. The generative process has *constrained* the state space of the canvas to a single possible state. It is known that $P(s_g) = 1$, and hence that the entropy of the system, $H(s)$ is 0, *if* the observer knows all about G.

A generative artist however in practice, even if designing the algorithm G, may yet not know what the outcome will be. They may know that the system will end up in *a* state, as they will know that this algorithm is deterministic, but they may not know prior to 'pressing play' which of the states it will be in. This will be dependant on how much effort the artist may have put into running the algorithm 'in their head'. In this case, G is a relatively straight forward process, and it is conceivable to predict the outcome prior to running the algorithm, if not exactly, they might have some intuition of the approximate *gestalt*, in particular if they are familiar or experienced with the generative process. However as the complexity of the generative process increases, it is

[6]and is 0 once they have observed it

increasingly difficult for a composer or designer to predict the outcome.

Nonetheless, knowledge of G, even if imperfect, will enable the artist/observer to make *some* predictions about the likelihood of certain states over others. It is trivial to know, for instance, that the starting state, where the canvas is completely white, is going to have probability 0 after running G. As such the probabilities distribution of S changes as knowledge of G is gained. In this case it, the entropy $H(S)$, will then be somewhere between the maximal value of the uniform distribution, and 0.

Knowledge of G, in information theory terms is the result of an *observation*. The amount of information I received from the observation, is equal to the degree to which uncertainty is reduced:

$$I = H_{before} - H_{after} \tag{2.2}$$

A reduction in H is a gain in information, but can also be cast as applying a *constraint*, so as to restrict the freedom of the 'choosing' of a state.

Even though the resulting artefact in fig. 2.2 is static, this final state however was reached via a number of intermediate steps, as each of the eight squares was added to the canvas. All processes happen *in time*. Each loop of G, as shown in fig. 2.1, causes S to *transition* from one state s_t to another s_{t+1}. It just so happens that G has the property of yielding perfectly cyclic pattern over 8 iterations, such that for $t > 8$, all s_t are equal; S enters stable static state. However that is not the general case, a generative process may, or may not have a defined end.

G itself can be understood as its own *system*, with three states g_1, g_2, g_3, which determines what instruction to carry out next: g_1 - draw a square, g_2 - translate, g_3 - rotate. The transitions from on state to the next is completely deterministic - it is known that state g_3 will follow state g_2, and that state g_1 will follow state g_3. The probability of the system at any one point in time being in a state $P(g_x) = 1/3$. However *given* that the previous state is known the probability becomes 1. For example given that one has observed g_2, the probability that state g_3 is 1: $P(g_3|g_1) = 1$.

It follows that not knowing what the state the system is currently in, the entropy $H(G)$ is maximal (1.58 bits). However once an observation is made on G, it is known what state it is in, and so $H(G)$ becomes 0. It is a deterministic process that will always yield the same outcome.

G thus is a system in its own right, with its own state space. It is however coupled to the canvas S, and the order and determinism defined in G is transferred to S. This is because the state

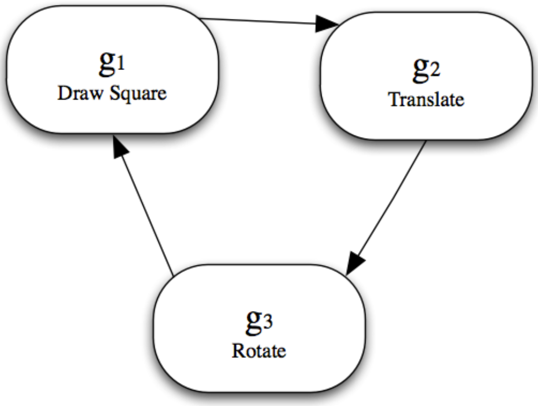


Figure 2.4: States of G.

of S is dependant on G. Without knowledge of G, the likely hood of guessing the final output state of S, s_g, is possible, just very very unlikely.

It becomes clear already with such a simple example above, that what constitutes a *system* is not always clear cut. Are the canvas, and the generative process two different systems, with one acting on the other? (G could be realised as a robot operating in the world, dripping paint on a literal canvas on the floor.) Or are they better consider as one system?

Or further should one consider the composer, who defined G to be part of the picture? There is no correct answer. In the same way that an output artefact can be viewed at different levels of description, as discussed in the previous section, any one particular system or sets of interacting systems can be broken down in numerous ways. What constitutes the system and its environment, where that boundary lies - and its ultimate subjectivity - is a topic that will be re-visited in our discussion of self-organisation in section 2.5.

2.3.2 Parameterisation

It is possible to describe a generative process partially, leaving some aspects of the process *underspecified*. One way to do this is with *parametrisation*. One can revise generative process G, such that the amount of translation in step 2, and the angle of rotation in step 3 are now not defined, and need to be specified as *parameters*, x and y. $G_{x,y}$ -

1. draw a square
2. shift the canvas x amount to the left
3. rotate the canvas by y degrees.
4. go to 1

Given that x and y have not been decided yet, it is not known what the final state of the canvas S, will be. However, knowledge of $G_{x,y}$, despite being underspecified, still greatly reduces the uncertainty over what the possible outputs may be.

In fig. 2.5 are a number of examples generated by $G_{x,y}$. As is readily apparent, the resultant outputs all seem to be 'of a class'. Despite being unique, they posses some perceptual similarities, in this case all being rather circular, with some having spiral or star-like properties. A designer who uses $G_{x,y}$, is, apart from some corner cases, doomed to make images of this class. On the flip side the amount of choice has been greatly reduced, from the near infinite dimensions of state space of the canvas - S - down to just selecting values for x and y.

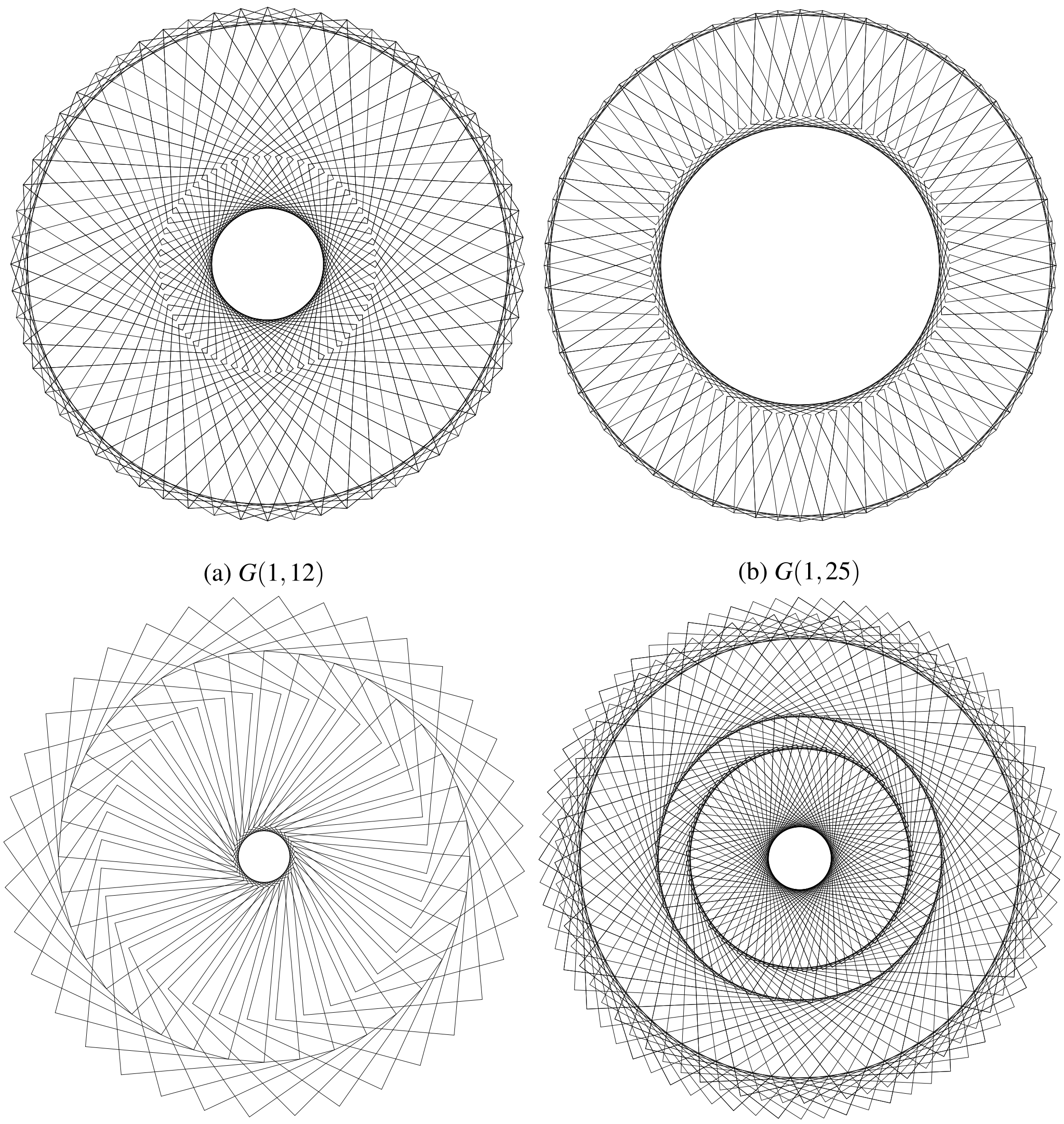
The space of all possible input parameters, is known as the *parameter space* of G: Θ_G. The process of finding the desired parameters is a process of *parameter search*. An outcome of a parameter search is denoted as finding an instance $\theta \in \Theta$. The space of all possible designs that can be generated defines the *design space* that can be worked with.

If the design space is now greatly constrained, the opposite, more positive view, is to consider these designs as having become *discoverable*. It would be much more difficult to find these designs without applying such formal procedures. In other words, the generative process *affords* the discovery of certain kinds of forms, that would not have been practically discoverable otherwise.

Given $G_{x,y}$, the uncertainty of the state of the canvas at a particular point in time, s_t, is now constrained to the entropy of the parameter space - $H(\Theta_G)$. As this is a deterministic process, once parameters have been selected, then the uncertainty drops to zero: $H(s_t|\theta_G) = 0$.

One can loosen the constraints of a generative process, increasing the size of the design space, by further under specifying the process. For instance instead of defining step 1 as 'draw a square', it could be re-cast as 'draw a rectangle with width w and height h'. The design is now significantly less constrained, however the parameter space Θ is orders of magnitude greater, correspondingly the uncertainty over the outcome prior to parameter selection, $H(\Theta_G)$ increases exponentially with each additional parameter.

The greater the complexity of a generative processes, generally the larger the parameter space that needs to be explored by the designer or composer. Beyond simple systems, parameter search is a non-trivial task. In some engineering and scientific contexts, parameter search can be cast as an optimisation problem, where a particular set of parameters can be evaluated against a measurable *fitness function* (or inversely a *cost function*). However when it comes to design or compo-



(a) $G(1,12)$ (b) $G(1,25)$

(c) $G(0.2,10)$ (d) $G(0.5,25)$

Figure 2.5: Instances of $G_{x,y}$ (100 iterations). x is the amount of translation relative to the unit length of the edge of the square, y is the angle.

sition, there is, generally, no such computable measure of the value of a set of parameters. There have been many attempts at defining such fitness functions, some of these will be discussed in section 2.6.3.

2.3.3 Stochastic processes

Another way to 'under-specify' a generative process, is to introduce non-determinism. As discussed in the surveys of generative art and music in section 2.2, stochastic techniques have been used by numerous artists and composers, particularly since computers became available.

As an example revisiting G, instead of having step 2 always follow step 1, and step 3 always follow step 2, and so on, the transitions *between* the steps, the states, can be made non-deterministic. Consider generative process $G_{stochastic}$:

1. draw a square.
2. shift the canvas unit length to the left.
3. rotate the canvas 45 degrees
4. go to step 1 with probability 0.5. go to step 3 with probability 0.5.

A number of possible runs of this process are displayed in fig. 2.6. As can bee seen, repeated runs of process $G_{stochastic}$ yields different outputs each time. Yet the works are nonetheless considerably similar; they are still just made from squares, all have this rather jagged somewhat disorganised, grid-like structures.

The corresponding state transition diagram is shown in fig. 2.7.

Such a stochastic process can be modelled as a *Markov chain*. Markov chains define the probabilities of a system transitioning from one state to another; for each state s_i, the probability of transitioning to another state s_j is defined in a *transition matrix*, M. Formally $P(sj|si) = Mij \in [0,1]$. The transition matrix for the above example is provided in table 2.1.

	g_1	g_2	g_3
g_1	0	1	0
g_2	0	0	1
g_3	0.5	0	0.5

Table 2.1: The Markov transition matrix for $G_{stochastic}$. The row corresponds to the current state, and each column contains the probabilities of transitioning to the next states. For example, given current state g_3, the probability of going to g_1 is 0.5, the probability of going to g_2 is 0, and the probability of staying in state g_3 is 0.5.

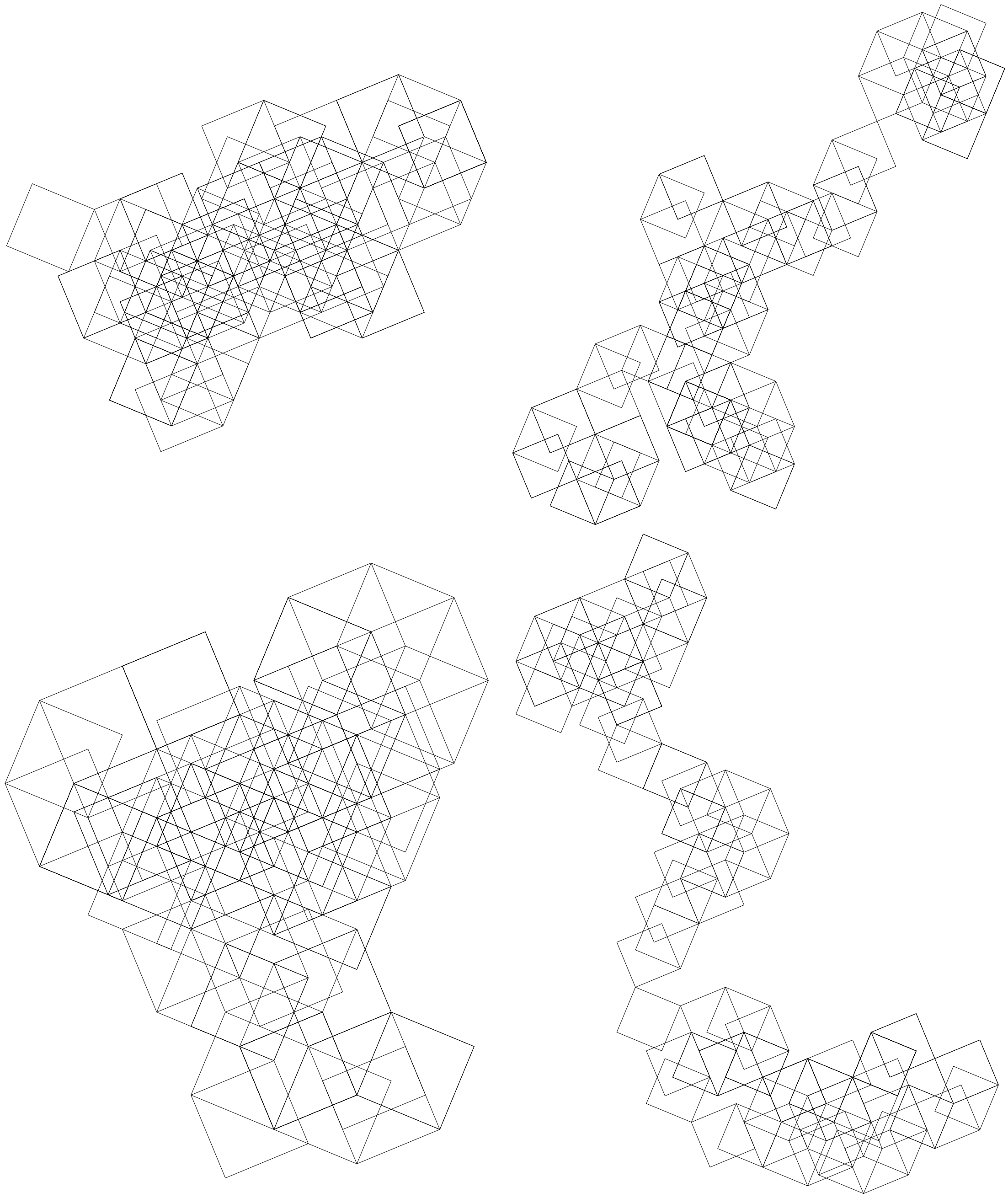


Figure 2.6: Instances of $G_{stochastic}$ (100 iterations).

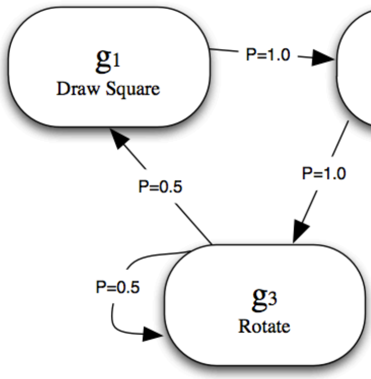


Figure 2.7: States and transition probabilities of $G_{stochastic}$.

Given a probability distribution $P(s_i,t)$ for the initial state, the transition matrix M defines the probability distribution for the next state: $P(s_j,t+1) = \sum MijP(s_i,t)$

Note how the values for transitioning from one state to another are asymmetric in time. The probability of going from $s_i \rightarrow s_j$, is not usually the same in the opposite direction $s_i \leftarrow s_j$. For instance one always shifts the canvas to the left after having drawn a square, but one never draws a square directly after shifting the canvas.

Given $G_{stochastic}$, there will always remain uncertainty over the outcome of the canvas; a different output would result each time the process is run. However there are *degrees* of uncertainty. This can be illustrated by modifying the generative process to have different probability in the final step. If one defines $G_{stochastic2}$ to be the same as $G_{stochastic}$, but the transition from g_3 to g_1 is more likely ($P(g_1|g_4) = 0.95$) and the transition to stay in g_3 is lower ($P(g_3|g_3) = 0.05$), then the output of $G_{stochastic2}$ has less uncertainty, as can be seen in fig. 2.8.

As can be seen, the output seems to be less disordered, yet are still fundamentally of the same style as each other. Underspecifiying a generative process through stochasticity allows a certain amount of surprise in the designs. The greater the amount of stochasticity, the greater the gamut of possible outputs, yet the less control, and predictability the designer has in the final design; the entropy of $G_{stochastic}$ is greater than that of $G_{stochastic2}$[7].

Naturally the parameterisation described in section 2.3.2 can be applied here as well; the probabilities of transitioning from one state to another can themselves be considered parameters.

The above examples only considered the previous state, as such they are known as 'first-order' Markov processes. It is also possible define the probabilities of transition based on a

[7]Additionally, our original generative process G can also be said to have a Markov transition matrix defining transitions between its steps. As it is deterministic, for any one state s_i there exists one state s_j such that $P(s_j|s_i) = 1$, however all other transition probabilities are 0.

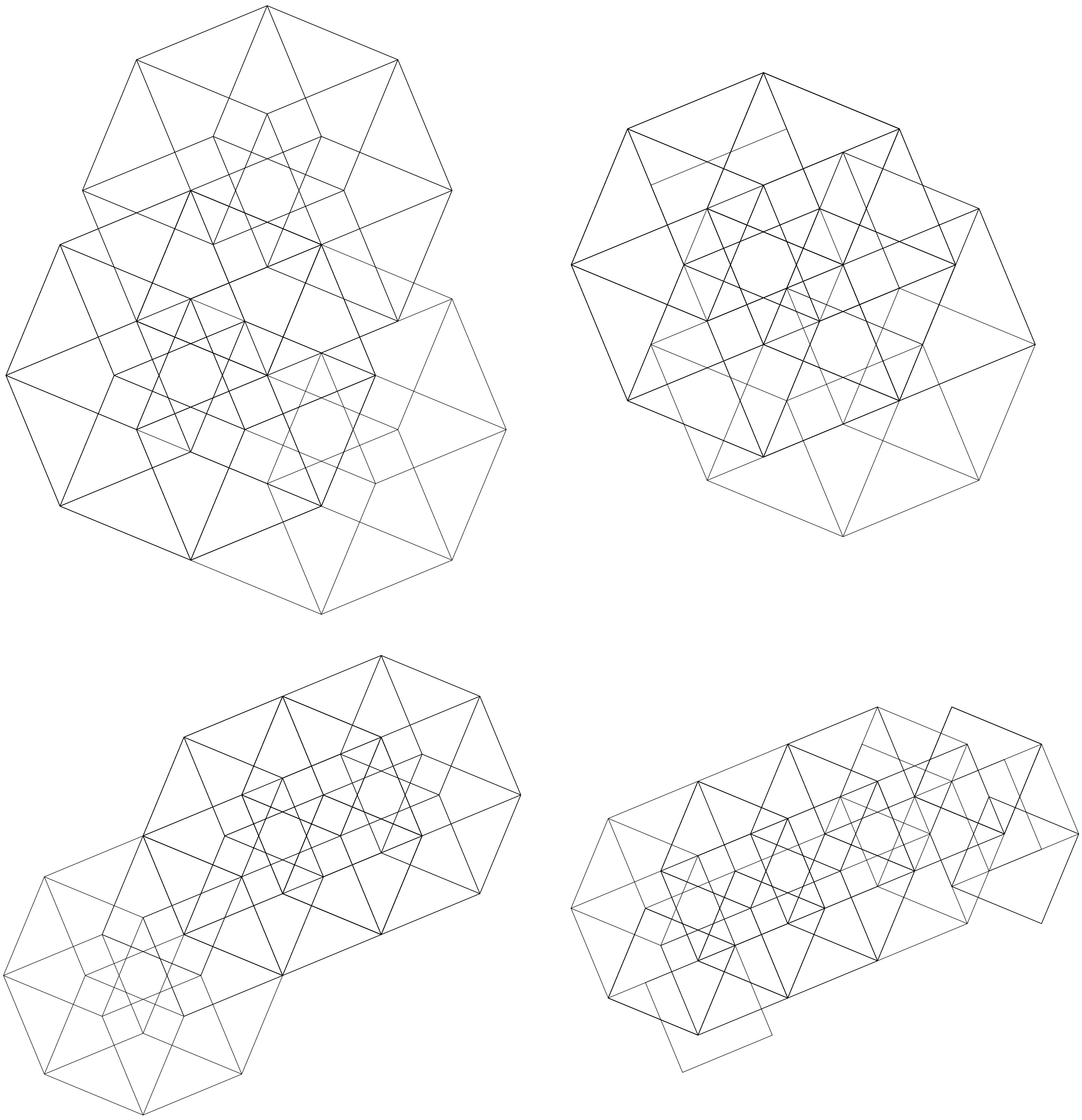


Figure 2.8: Instances of $G_{stochastic2}$ (100 iterations)

longer history of transitions. Such a 'higher order' Markov chain might state that given states s_i and s_j have happened, define probabilities of going to the next state. This would be a 2nd-order Markov process. This can naturally be extended indefinitely to Nth order, such processes are also known as *n-gram* models. Higher order Markov process however can always be reduced to a first order transition matrix by grouping states in to meta-states, each representing a sequence of n states.

As discussed, Markov processes have been used extensively as generative mechanisms both in computer art and in music. Markov chains are useful for modelling stochastic processes, and will be explored extensively in the studies of this thesis; the *Melody Triangle* is an interface for exploring the musical possibilities of Markov processes, while *EvoColour* uses 3rd-order Markov processes to generate images evolved in an interactive evolutionary algorithm based on selections by the public. Markov transition matrices are a way of representing uncertainties, and how much uncertainty, or conversely predictability, is preferred by viewers or listeners of a process is one of the questions addressed by the studies. Markov chains will be revisited in the chapter 3, where a number of information measures that quantify this uncertainty will be discussed.

2.4 Contemplating the Output

Having outlined some of the basics and peculiarities of generative processes, this section considers issues regarding the *consumption* of the generated artefacts by observers.

Notions of similarity will be discussed with respect to state spaces, and then with respect to the parameter space of generative processes. Similarity, with regards to features and feature spaces, are then outlined. This leads on to Peter Gärdenfors theory of *conceptual spaces*(Gärdenfors, 2004b); a framework that supports reasoning in geometrical terms, *quality dimensions* defining perceptual similarity with geometrical terms such as 'betweenness' and 'closeness'.

A discussion of value and aesthetic theories, will be provided in section 2.6.

2.4.1 Similarity

If a completely stochastic process is applied to a melody, say by rolling a dice for each note, then the output would be completely random. If this process is repeated, a different melody will be generated, different from the last. Such completely random melodies, even though unique and completely different from each other, from the point of view of a human perceiver will be comprehended as essentially the same – they will just be heard and understood as random melodies, with no discernible structure or patterns.

As such, in some important way, they have an equivalence, despite their uniqueness. Similarly if one picks two random states from S, the state space of the 100x100 binary display, it is very likely that the selected images would be two different instances of something that is comprehended as just 'noise' (see fig. 2.9) Although in terms of the machine representation, the bits in memory that represent the pixels, they are completely unique.

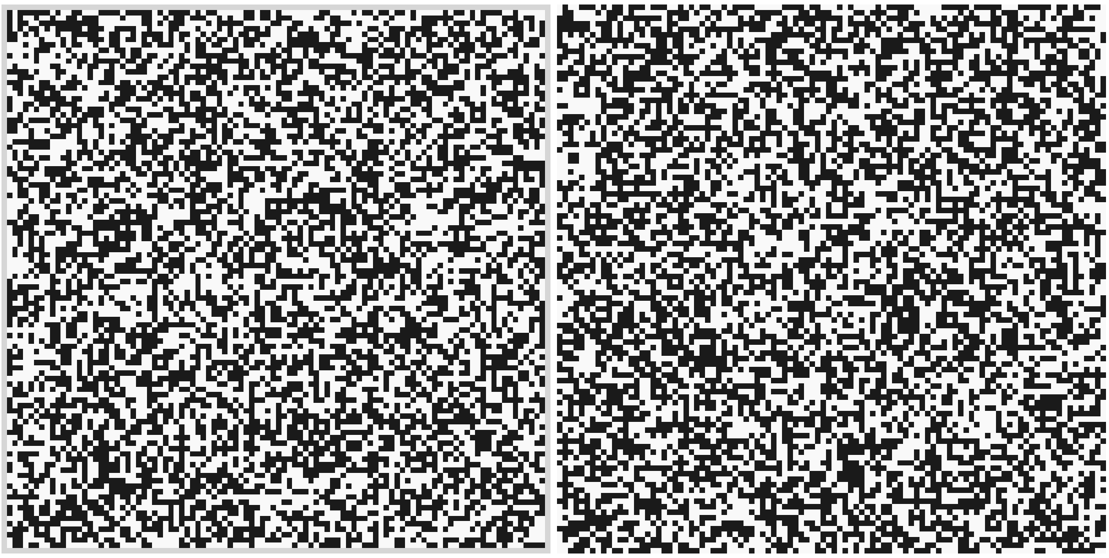


Figure 2.9: Two random states from S, 100x100 binary display

Each state is represented by a number of variable features. In the case of the 100x100 binary display, that would consist of 10000 binary variables. That is $s = v_{1,1}, v_{1,2}, \ldots, v_{100,100}$. Two states that are 'far apart', in terms of the euclidean distance of the set of low-level variables that define S, can still be *perceived* as very similar. However conversely, it is likely that two states that differ from each by one variable, say modifying one pixel in our display, will be also perceived as very similar. In short, there is little that the distance of the state's variables the tells us about similarity with certainty. In many cases, states that differ very slightly in terms of variables, will look very much the same, but that may not even always be the case, depending on the level of description at which the system is being considered.

Considering the generative process $G_{x,y}$ (the parametric version of G defined in section 2.3.2), and defining three instances of it g_1,g_2 and g_3 (the first with values x_1 and y_1, the second with x_2 and y_2, and the third with x_3 and y_3), if x_1, x_2 and x_3 are set to be the same, a unit distance, 1, and y_1, y_2 and y_3 are set to differ only *slightly* from each other (45, 46, and 43 degrees respectively), the outputs are very different, as shown in fig. 2.10

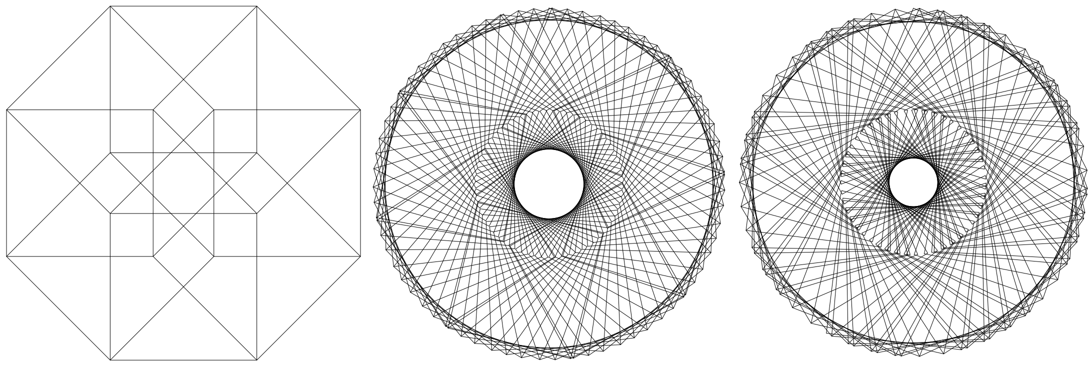


Figure 2.10: Three instances of $G_{parametric}$ - g_1,g_2 and g_3, with the angle parameters set to 45, 46 and 43 respectively.

As can be seen, both g_2 and g_3 are very different from g_1. However g_2 and g_3 are more visually similar. As such, it is clear that for a generative processes, a small 'parametric distance'[8] between two sets of parameters, may, or may not yield perceptually similar outputs. There is little that can be predicted with regards to similarity by looking at parameter values, especially as systems become more complex.

Therein lies some of the difficulties faced by algorithmic composers and generative designers. Having defined their generative processes, they engage in an exploration and experimentation of the *design space* within which they work in the form of parameter search and discovery. However in a generative processes, the causal link between the selection of input parameters and the generated output can be non-linear and difficult to predict, even with simply generative algorithms as demonstrated above, due to the emergence that occurs as low-level elements interact or intersect, bringing forth higher-level patterns and forms.

The situation becomes particularly salient when systems are complex and the output is determined by numerous inter-connected components. In these systems, the externally perceivable properties of the system become increasingly difficult to be deduced from its components. This makes the parameter selection process, and finding what are 'good' parameters, a difficult problem for designers and composers.

In 'explicit' algorithmic creation there are two primary consideration - the design of the generative system and its parameterisation, and the selection of parameter values. Artistic form

[8]Naturally there a different ways of defining a 'distance', one could use a 'city-block' metric, or euclidean distance for instance. Moreover, some parameters may be defined differently from each other; some may be real numbers, some binary values, and so a measure of distance becomes more difficult to define in those cases. However, the overall point remains the same.

emerges out of the artist's interaction with her medium; the dynamics of the feedback loops between the artist, the abstract system and its output is crucial in determining how naturally and intuitively the aesthetic potential of an algorithm can be discovered. Therefore specific attention must be paid to the task environment where the compositional/design process takes places, in particular to the design and implementations of user interfaces.

The mapping from parameter space to phenomenal output is of particular concern to the designers of tools and instruments, where the generation happens in real-time. Considering the many works in the field of New Interfaces for Musical Expression (NIME) for instance, here great effort is spent in mapping gesture to engender and facilitate intuitive expression and musical performances.

However in generative art and music more broadly, there is not necessarily this consideration for 'intuitive' mapping from parameter space to perceptual properties of the generate output. Consequently exploring the possibilities afforded by generative processes often ends up being done in a fairly ad-hoc fashion. Moreover design involves exploration and discovery; designers may not know *apriori* what the features of a desired output are.

Part of the generative artist's creative process can thus be characterised as one of *navigation* or search through the possible space of input parameters. Such a search through parameter-space commonly consists of discrete steps of parameter selection and output evaluations that feedback into the design of the form-generation processes. In contrast to more traditional methods of creation where there is immediate and clear causality between action and output, in algorithmic creation the 'action-perception cycle' can be discontinuous.

Features

It is possible to make objective, numerical measurements on an output artefact. Such an individual measurable property is a *feature*. Feature measurements are central aspect of many statistical and machine learning techniques across disciplines, and an essential process in machine based pattern recognition and classification across mediums, be it sound, vision or analysis of raw numerical data.

Starting with a data set, such as an image, a process of *feature extraction* results in derived values that in some way describe the artefact under consideration. For instance, numerical features on an output from G, could be simply, the 'number of white pixels', or the 'number of black pixels'. *Feature construction* is the result of defining 'higher-level' features as a combination of other already extracted features, this could be for instance ‘the ratio of white pixels to black pixels’.

In an analysis tasks, features can be combined to construct a *feature vector* to represent an instance of an object, the corresponding space of possible features is the *feature space*. Any one feature describes one objective measure, and depending on the task at hand, be it classification or identifying similarity, a feature may or may not be relevant, or may be redundant with respect to other features in the vector. A feature may, or may not, have correlation with how a human comprehends an output, or have any correlation with similarity.

A feature such as ‘the ratio of white pixels to black pixels’, f_{bw}, correlates well with perceptual similarity in our examples of g_1, g_2, and g_3 - the value for g_1, ,is further away then the values for g_2 and g_3, which are quite close to each other. However this is only with respect to G. If one attempted to define similarity with this feature with respect to the more general state space the canvas S, there are many cases where this no longer holds, as examples below in fig. 2.11 demonstrate. Those two images have the same f_{bw}, yet are completely different from each other perceptually.

Figure 2.11: Two states of S that have the same value for f_{bw} (0.5).

As such, the relevance of a feature is dependent on both the task at hand (e.g. measuring similarity) and the space of possibilities under consideration (e.g. the state space of all binary images S, verses the state space of G). In the case of image analysis, more refined features that take into consideration the spatial properties of an image need to be taken into consideration. This can include feature extraction processes such as edge detection or blob detection. The ‘task at hand', for instance reading number-plates on cars, or recognising faces, will strongly shape the choices of features considered.

Features do not 'exist' in a system in any real sense, they are found and defined by an observer to describe a situation numerically and objectively. As such it follows that any one particular feature space, may or may not be effective in describing a situation. Further in the same way that the mapping between parameter space to perceptual similarity is non-linear, the mapping from parameter space to any one particular feature space may be non-linear and difficult to predict as well.

However if a feature-space is well defined, computational techniques can yield insights in analysis and help correlate the features to perceptual mechanisms. Further they can also be used to define new creative generative processes, or provide a framework efficient means of parameter space navigation. The *Melody Triangle* described in chapter 3, attempts to provide parameter space navigation in terms of features related to perceived predictability.

For effective statistical analyses, often the feature space needs to be pruned of redundant and noisy features. To this end a number of techniques, such as *principal component analysis*, can be employed to reduce the dimensionality of the feature space, applying a weighting to the individual features to reflect their relevance to the task at hand. This is a process of *feature selection*.

Which features are relevant to the aesthetic evaluation of a system are dependent on, amongst other things, the faculties of the observer. For instance a completely colour blind individual would not perceive any of the features that are colour related, but would still consider other features such as form and shape. Similarly it is possible to define a set features that have no relevance to the aesthetic evaluation, or the phenomenal perception of the work.

2.4.2 Conceptual Spaces

As has been discussed in the previous sections, quantifying similarity is a non-trivial process when looking at states spaces, parameter space, or even feature spaces. Peter Gärdenfors *Theory of Conceptual Spaces*(Gärdenfors, 2004b) is a framework for representing concepts based on geometrical structures. Within conceptual spaces, similarity can be modelled in a natural way by relating similarity to proximity in the geometrical structures. Gärdenfors argues that judgments of similarity are key to understanding a large number of cognitive processes: "judgments of similarity reveal the dimensions of our perceptions and their structures"(Gärdenfors, 2004a).

Quality Dimensions

A number of *quality dimensions* together define a *conceptual space*. Each quality dimension represents the various 'qualities' of an object. Within a conceptual space, proximity and distances relate to similarity, and conversely, the further apart two objects are, the more dissimilar they are. The dimensions form the means by which properties are assigned to objects, and serve as framework for identifying relations between them.

A quality dimension can correspond closely to how sensory receptors receive information from the world. For instance, brightness, temperature, or pitch are quality dimensions that are related to basic sensory processes. However quality dimensions can also be of "an abstract, non-sensory character"(Gärdenfors, 2004a), and further new quality dimensions can be learned. He cites Piaget's famous experiments, where small children are unable to distinguish between height and volume, as an example of an expansion of quality dimensions. Gärdenfors further suggests how some quality dimensions are culturally acquired, using the example of time as being conceived as circular in some societies to support this view.

Finally he suggest that quality dimensions can be discovered by science, and learned by students of science. The distinction between the quality of dimensions of mass and weight, as proposed by Newtonian science, is an example: "Learning new concepts is ... often connected with expanding one's conceptual space with new quality dimensions"(Gärdenfors, 2004a). It is reasonable to suggest that the learning of conventions and norms of an artistic or musical style could also be understood as an expansion of conceptual spaces.

Concepts

The suggested epistemological value of conceptual spaces is to "serve as a tool in sorting out various relations between perceptions"(Gärdenfors, 2004a). Formally a conceptual space S_c[9] consists of a set of n quality dimensions $D_1, \ldots, D_n$. A point in conceptual space is represented by a vector with one value for each quality dimension - $s_c = < d_1, \ldots, d_n >$. Two objects are located close to each other in this space, in Euclidean terms, if, and only if, they are judged to be *similar*, and the further apart they are, the more dissimilar they are.

Gärdenfors proposes the idea of a *natural concept* with the following criterion: *"A natural concept is a convex region of a conceptual space"*(Gärdenfors, 2004a). From this it follows that any two points within a such a convex region, will necessarily belong to the same natural concept.

[9] the subscript c used here to distinguish between the *state space* S described earlier

He conjectures for instance that natural colour terms, such as ‘red’ and ‘blue’ will correspond to natural concepts in the conceptual space of colour. That is, if an object can be described with a particular natural concept, such as the colour *red*, and another second object is also deemed to be red, then any other object that has a colour in between the first two objects, in the geometric space defined by the quality dimensions, must also be assignable to the colour red, because of the convex shape of the natural concept.

However it must be understood that this extrapolation can *only* be made if the quality dimensions, and hence the conceptual space, is well defined with respect to human perception. Not all representations of colour space have a correspondence with quality dimensions. Colour theory and the developments of colour models will be discussed in section 5.2 of the *EvoColour* chapter. However it is worth briefly mentioning that for Gärdenfors the discovery of the colour-circle, that is the understanding of hue as a cylindrical coordinates, is a kind of ‘discovery’. With the dimensions of lightness and chroma, as discovered by Munsell (and discussed in later sections), they together define a conceptual space, allowing for geometric notions of similarity and betweenness, as well containing convex regions for natural concepts.

This contrasts with the RGB colour space. RGB is aligned to how colours are generated by a computer monitors in an additive model, which does not carry inbuilt considerations of how the human visual system processes colours. In this space, betweenness and parametric proximity, do not necessarily align well with perceptual similarity, nor are there clearly defined convex regions that form natural concepts, as they do in the Munsell system (or modern manifestations thereof, such as HSL or CIELAB).

In the same way that the Newtonian model introduced the distinction between weight and mass, understanding that some aspects of human perception of colour is more accurately described with a cylindrical geometry, supports not only better reasoning about colour, but affords the possibility of designing systems to allow designers to more easily choose and navigate and find colours.

As the research into colour perception has progressed, more refined models have been developed, many iterations beyond the Munsell model. The most current colour models used in colour perception research, such as the CIELAB colour model, have further considerations to make the euclidan distance between colour values in the space correspond more accurately to experimentally derived measures of human judgments of similarity. It then follows naturally that the most sophisticated tools for designers, such as the more recent versions of Photoshop, provide tools for designers to manipulate colours with respect to the CIELAB colour space.

Quality Dimensions in Generative Processes

The conceptual spaces framework can serve two purposes, one is as an explanatory framework to help in reasoning about cognitive processes. The other is that it can provide a framework for the construction of artefacts that carry considerations for human perception.

Within any particular domain, or any one particular generative system, it is not clear exactly what the quality dimensions are, or even necessarily how many such quality dimensions there are. A conceptual space is well defined if human judgments of similarity correspond to proximity in the quality dimensions. In earlier sections, it was illustrated how the notion of similarity is elusive when it comes to generative processes; similarity is not easily found in state spaces, parameter space or even feature space.

However, if quality dimensions for a generative process can be identified, then it follows that a generative artist may be able to design into systems and their interfaces, the means of navigating the systems' parameter space in terms of similarity of output. This can lead to more effective reasoning about the processes, and ultimately faster discovery of parameters of values; in the same way that reasoning about, and searching for colours in HSL is more intuitive then in RGB. In a generative system, this could be done by re-aligning parameter space to map to correspond to quality dimensions of the outputs.

Further, if one is to reason about *aesthetic value* in generative artefacts, doing so in terms of the quality dimensions of a conceptual space would provide more effective grounding and descriptive language within which to do this reasoning, then, say, attempting to do this reasoning in terms of the state spaces or the parameter values directly.

Quality dimensions can be defined from feature spaces. A quality dimension could correlate to a feature directly, or could be defined as weightings and combinations of various features. However identifying the salient features and feature combinations is a challenge. One approach is to conjecture the relevant quality dimensions around some theoretical frameworks and models of human perception, a bottom up approach. If the resulting geometrical topologies yields results such that perceptual similarity follows proximity in the geometric space, then it would suggest that perhaps some aspect of the conceptual space, for the context at hand, has been uncovered. Further, if convex areas in the space have defining characteristics, it is possible that these areas are representing natural concepts.

Another approach, top-down, is to arrive at the quality dimensions through measurements of similarity in experimental contexts. The most sophisticated colour models, such as CIELAB, were arrived at and iterated upon over many decades of experimental evidence and measurements on colour relationships and similarity. This model will no doubt be further refined and improved upon, and incrementally represent the conceptual space of colour more accurately.

A well defined conceptual space can predict similarity. However when it comes to predictions of judgments of aesthetic value, the situation is considerably more diffuse. As elusive as quantifiable aesthetic value may be, any such attempts will be best positioned to reach its goals if models that align with conceptual space are used. This is reflected in the way that contemporary research in colour preferences and harmony have been using incrementally more sophisticated colour models, with CIELAB the current standard practice.

It follows that judgments of aesthetic value on artefacts (assuming that value is not *entirely* 'in the eye of the beholder'), can hint at an underlying orderings and organisation, and further, can perhaps help uncover some of the geometrical attributes of the relevant conceptual spaces, which in turn shed light on the underlying perceptual mechanisms.

Conceptual spaces will be re-visited throughout this thesis. In chapter 3, the *Melody Triangle* is an interface derived from an information theoretical approach to modelling human listening. Information measures that demarcate a geometrical space that aligns well with perceptual similarity define a musical interface. These measures hints at a conceptual space of predictability, for a very restricted aspect of the musical listening experience. In *EvoColour* (chapter 5), considerations of similarity relationships in the conceptual space of colour were inbuilt into the algorithms of an evolutionary system. Further, the evolutions, driven by judgements of human preference, begin to hint at underlying orderings and relationships on the attributes of these images. The resulting analysis suggesting some aspects of a possible geometry of an underlying conceptual space.

2.5 Self-Organisation

Whence does the appeal for generative processes and the resultant emergence arise? As already pointed out in section 2.2, there is ubiquity of the appeal of patterns and generative processes across cultures and across media. One can conjecture a mimetic view of aesthetics, whereby artists reflect upon processes in nature, could be the source of the appeal, for nowhere is emergence more pronounced then in nature.

However there is one crucial difference between the emergent forms that are ubiquitous to nature over those that are simply drawn or composed (algorithmically or not) by a creator - that of *self-organisation*. Religious considerations aside, the forms and structures and patterns in nature seem to have not been put in place by an external set of instructions that are somehow followed, there is no external '*G*' that has placed order in the world.

Many of the forms and structures that are seen in nature can be emulated with algorithms; swarm simulations can show accurate behaviours of flocks and shoals, L-systems are remarkably effective at describing the structures of plants and trees, such processes explicitly the per-view of artificial-life systems. Yet there is no sense in which explicit instructions, like those in *G*, 'exist' anywhere in nature. In nature, structures, be it relatively simple ones such as minerals, straight through to complex ones like the human brain, seem to be a product of some sort of 'self-organisation'.

The theory of self-organisation has grown out of variety of disciplines, including thermodynamics, and systems science and cybernetics. A brief overview of the basic concepts of self-organisation will be presented, as they relate not only to the designs of many types of processes designed by generative artists, but also further the studies in this thesis can all be understood in some sense with regards to these concepts. Further, some of the concepts of self-organisation can be used to reason about the studies *in situ*, by considering the human contemplators as part of an overall process, the data collection itself as a process of organisation.

Heyligen defines self-organisation "as the spontaneous creation of a globally coherent pattern out of local interactions"(Heylighen, 2001). The cybernetician W. Ross Ashby in formulating his "principle of self-organisation", suggested the dynamical systems tend to always reach a state of equilibrium. As such a state reduces the observers uncertainty about the state of the system, it can be measured as reduction of entropy(Ashby, 1991).

However, an intuitive view of this would suggest that there is a disconnect between self-organisation with the Second Law of Thermodynamics, which states that entropy, the disorder in a system can only increase, and not decrease. How then could it ever be possible to have self-organisation? The answer to this seeming impossibility comes from understanding that self-organising systems cannot be understood in isolation, they are part of a wider environment with which there is a flow and exchange of matter and energy. For instance, living entities take in low-entropy in the form of food or sunlight, and dissipate high-entropy waste into the environment. As such, it is just as valid to consider self-organising systems as 'dis-organizing systems'(Von Foerster, 2003b) as they 'disorganise' their surroundings.

2.5.1 Quantifying Self-Organisation

In his text, 'On Self-Organizing Systems and Their Environments', the cybernetician Heinz von Foerster draws a distinction[10] between a system S and an environment E. He describes this situation with reference to the Second Law of Thermodynamics, and states

> "the system is in close contact with an environment, which posseses available energy and order, and with which our system is in a state of perpetual interaction, such that it somehow manages to 'live' on the expenses of this environment"(Von Foerster, 2003b).

To conform to the intuitive definition of a self-organising systems, it follows that self-organising systems should be "increasing their internal order" over time. Von Foerster proposes Shannon's information measure of 'relative entropy', also known as *redundancy* as a way of quantifying the amount of order in a system -

$$R = 1 - H/H_m \tag{2.3}$$

Whereby H is the entropy of an information source, and H_m is the maximum possible entropy that the system can have. As such if $H = H_m$, the system is at its maximum state of disorder, and R is thus 0. Conversely, if the system is in a state of perfect order, $H = 0$ and then redundancy becomes 1.

Given the measure of *redundancy* as a measure of order, it follows that for a self-organising system, the *rate of change*, δ, of redundancy, R, should increase over time:

$$\delta R/\delta t > 0 \tag{2.4}$$

[10]The notion of a distinction, of a border between systems and environments is understood from a cybernetic perspective as being a slippery, and ultimately subjective concept. This idea has already been touched upon earlier, and will recur throughout this thesis.

This means that as time progresses, the redundancy of the system increases. Differentiating with respect to time, and using the definition of R, von Foerster provides the condition for a system to be self-organising expressed in terms of entropies:

$$H(\delta H_m/\delta t) > H_m(\delta H/\delta t) \tag{2.5}$$

It follows from this, that there are two 'corner cases' whereby a system can be self-organising. The first is the case where the H_m is constant, that is the maximal entropy of the system cannot increase. It follows from this that - $\delta H/\delta t < 0$, that is, the entropy of the system must decrease over time, corresponding to the intuitive understanding of a self organising system.

The other corner case however is more subtle. That is, when H is constant - the current entropy of the system stays the same - it follows that $\delta Hm/\delta t > 0$. This states that a system can be 'self-organising' if the maximum entropy of the system increases, while the current entropy remains constant.

One way to imagine this, is if a system is being 'added to' but in such a way that order maintained. For instance imagine collecting cards to form a deck. If as each card is added to the deck, it is placed not just on top, or in a random position in the deck, but is placed in its corresponding place in numerical order. When a card is being added, the H_m increases, as the maximum possible disorder increases as the deck size increases, but the current H of the deck remains minimal.

It follows from this the counter intuitive notion that there can be self organisation *even when* the entropy of the system is increasing: 'self-organisation' occurs as long as the maximum possible entropy H_m increases *faster* than the rate of increase of the current entropy of the system H.

2.5.2 Data Collection as Self-Organisation

The experiments in this thesis can be considered 'systems' that interact with an 'environment'; the subjects of the experiments. Different types of data collection are present in the system, in some cases, as in the first corner case above, the maximal entropy of the system remains constant. For instance in the crowdsourced evolutionary algorithm, *EvoColour*, populations of images remain constant in size, and as the selective pressure of the users is applied to the populations, the internal order of the systems increases. In the experiments with the *Melody Triangle* mobile phone app, users submit their preferred settings, and these collected settings accumulate. For each submitted setting, the maximum possible disorder of the system, H_m, increases constantly. The current entropy of the system also increases, however as the user's preferences are (presumably) not completely random, the rate of increase of H will not increase *as fast*.

But how is entropy, the amount of disorder in a system to be calculated? This is very much a subjective consideration. To illustrate this, consider this classic example von Foerster uses throughout many of his lectures and papers(Von Foerster, 2007). Given these two numerical sequences, which is more ordered?

- A: 0,1,2,3,4,5,6,7,8,9
- B: 8,5,4,9,1,7,6,3,2,0

The natural view is consider sequence A as being more ordered, and hence of lower entropy, than sequence B. However this only holds, until it is pointed out that second sequence is perfectly ordered, but with respect to alphabetical order (*'eight', 'five', 'four', 'nine', 'one', 'seven', 'six', 'three', 'two', 'zero'*). This illustrates that the order, and hence measures of entropy, are subject to the framework and language with which this order is considered, and that changing the language by which order is considered, different orders are perceived and measured. Entropy and order are subjective. As von Foerster exclaims, "disorder and order are our inventions!"(Von Foerster, 2003a).

The subjects of the experiments presented in this thesis are made to, in some way, give input to a system in terms related to perceived aesthetic value; be it with preferred settings of a mobile app, gaze points from an eye-tracker, or explicit preference judgments between pairs of images. As such finding the features, or features combinations, that most efficiently encode the decrease in the overall system disorder, should indicate that the selected features in some way relate to the aesthetic value, for their particular contexts. Further as suggested in section 2.4.2, finding feature combinations that correlate well with aesthetic judgments, can yield insights into the underling conceptual spaces, with which it may be possible to predict what kinds of outputs may be deemed of value. Naturally there are a number of caveats. For instance, how well does the order that the systems extract from the users *really* correspond to aesthetic value, and are not, say, artefacts of interface priming. Another consideration is wether or not the features required to identify the increase in order can be found. These considerations will be re-visited in each of the studies.

2.6 Aesthetics

The next section considers aesthetic theory. As this is indeed a vast field, only a brief survey of key historical themes in aesthetic thought will be provided. This is followed by a discussion of empirical aesthetics, attempts to measures of aesthetic value, and information theory in aesthetics.

2.6.1 Philosophical foundations

Aesthetics was long the purview of philosophy, and not till more recent times have there been attempts at understanding the aesthetic experience from a scientific perspective. One of the core threads of the history of aesthetics is concerned with to what extent the beauty is *objective*, or *subjective* (the *objectivist* and *subjectivist* views respectively.)

Early views such as those proposed by Plato and Aristotle were largely objectivist, and considered art as *mimesis*. In a mimetic approach, value relates strongly to how well the object 'reflects' a reality. Plato suggest that there exists ideal and pure forms – *Platonic ideals* – and that artistic objects were as imperfect copies of these abstract pure forms (and consequently due to their imperfect nature, to be an inferior, and ultimately negative influences). An Aristotelian views similarly considers artistic objects as forms of mimesis, but more positively and that the delight they provide should be sought after(Graham et al., 2005).

The subjectivist view on the other hand proposes that aesthetic value is not a property of the object per-se, but are the result of the appreciating process of the perceiver, as per the famous saying 'beauty is in the eye of the beholder'. A number of philosophers, such as Kant and Hume have advocated that there are both subjectivist and objectivist elements to the perception of beauty. David Hume for instance states: "Beauty is no quality in things themselves: it exists merely in the mind which contemplates them; and each mind perceives a different beauty"(Hume, 1985). As such he supports the subjective view, all opinions are 'all right', because the value is not a property of the object itself. Despite this he also argued for universal standards by which judgments of beauty are made, and that the ability to make sound judgment is something that is learned through knowledge and training(Shimamura, 2012).

The history of western fine-art followed an initial trajectory towards more and more photo-realistic works, as increasingly sophisticated techniques (such as the understanding of perspective in the Renaissance) were applied. However this was then followed by a tendency *away* from mimesis. Romanticism placed greater emphasis emotional expressiveness, augmenting the mimetic with exaggerated emotional inducing content. Leading on to expressionism, the emphasis to communicate and induce emotions to the beholder, independent of the realistic forms of objects, as exemplified by Matisse's famous statement "I do not literally paint that table, but the emotion it produces upon me"(Matisse & Flam, 1995, p. 66).

With the advent of Modernism and Impressionism, the emphasis continued to shift further away from *mimesis* to the 'impressions' of nature, emphasising the sensual quality of paint and light itself; art would be appreciated on the basis of its sensory qualities - that is, the aesthetic interplay of colours, lines, textures and shapes. It has been suggested that this shift away from the mimetic was as a response to the advent of photography, as it rendered mimetic paintings as inadequate or outdated(Shimamura, 2012).

Finally with the advent of conceptual art and post-modernism as exemplified by Duchamp's famous *Fountain*, the concepts and statements about art became the subjects of art. Here the sensory qualities of the objects become a secondary, if not irrelevant concern, and value relates to the concepts addressed, and as such require prior knowledge of social context.

A succinct summary of philosophical considerations is provided by Shimamura:

> Beholders over the centuries have considered artworks in terms of (1) how successfully they mimic the sensory experience of looking at the real world as if through a window (2) how well they express feelings and a sense of beauty (3) how well they create significant form, and (4) how well they convey conceptual statements(Shimamura, 2012).

How then does generative art fit into this picture? Generative art, on the surface, seems not be 'mimetic' of sensory experience in a direct sense. As discussed, the forms and patterns made are often abstract, geometric, and not a direct pictorial representation of a scenes or objects 'out there'. Further it was discussed in previous sections that the motivations for the use generative techniques are widely varied(Galanter, 2003), ranging from the spiritual to the scientific, to simple convenience.

However, as generative processes are ubiquitous to nature, in another way it could be argued that there is a mimetic element to generative art, as a reflection of the processes of nature. Some of the works of Xenakis, such as those exploring the statistical behaviour of gasses for instance, come to mind. Evolutionary processes and a-life art generally also contain explicit mimetic considerations. Although these works may at times look like like pictures of nature (such as McCormack's L-system-based artificial plant growths), more often the forms generated are otherworldly and abstract. The mimesis is nonetheless there, but at the abstract level of the latent underlying processes.

McCormack and Dorin similarly draw a parallel between nature and generative processes. Considering Kant's concept of the *sublime* – the sense of wonder, and even terror, brought forth when contemplating the overwhelming boundlessness of nature – they suggest that a sensation of the sublime can manifest itself in the *emergence* of generative processes, particularly when mediated through the vast scale and speed of computation, a concept they call the *computational sublime*: "the instilling of simultaneous feelings of pleasure and fear in the viewer of a process realized in a computing machine"(McCormack & Dorin, 2001).

Additionally some generative art tackles conceptual considerations. Clear examples are the systems that themselves 'make' art or compositions, such as Cohen's famous *AARON*, or Cope's *Emmy*. Beyond the phenomenal quality of the works they create, these systems pose difficult philosophical questions relating to the nature of creativity itself(Boden & Edmonds, 2009), and some of these issues are discussed in section 2.6.5.

2.6.2 Empirical Aesthetics

Largely inspired by the *objectivist* view, aesthetics have been explored empirically. One of the earliest such attempts was in 1876 by Gustav Fechner. In his *Primer on Aesthetics*(Fechner, 1876) he suggested that aesthetics could be studied from a systematic bottom up approach, in contrast to considering complex philosophical concepts such as 'beauty' and the sublime.

In his experiments he studied preference judgments for basic shapes, such as rectangles in various proportions. The general thesis being that by understanding the basic elements of visual aesthetics, it would be possible to construct a general theory aesthetic experience. There then followed countless experiments that adopted this techniques, where observer are asked to rate shapes, colours, objects in terms of preference over others, under controlled conditions; an approach still practiced to this day.

Gestalt psychology proposed an alternative to this bottom-up approach, presenting a more holistic approach to perception. Their famous credo 'the whole is other than the sum of the parts' argues that perceptions cannot be dissected into basic elements. It proposes a number of 'laws' or 'principles' dealing with the modality of vision, that relate how low level stimuli are grouped or associated into higher level wholes. This includes the principle of *Prägnanz*, suggesting that perceptions are organised by the most succinct interpretations(Smith, 1988, p. 61). A Gestalt view correlates well with perceptual emergence that is visible in the generative processes. For instance as has been illustrated in our discussion of generative processes, one more readily perceives the higher level forms than the low level elements.

Arnheim in his book *Art and Visual Perception*(Arnheim, 1954) draws on Gestalt theory to offer a theoretically motivated approach to the psychology of art. He considers the way images are organised with respect to "perceptual forces" in terms of principles of organisation. He suggests that artists through balance, harmony and object placement manipulate such perceptual forces to yield aesthetic experience. His first example, of a circle inside a square, but placed to one side, is said to 'heighten tensions', where as if the circle is in the middle of the square, reduces tension.

Berlyne considered more directly the affects that artworks have on beholders. He suggests that there is an optimal amount of tension that yields the optimal 'hedonic value', in a perceiver(Berlyne, 1970). If an object is too boring and predictable, then the there is little pleasure to be had, however conversely if an artwork has too much tension, if it is so terribly arousing it can cause confusion or displeasure. He suggests that we prefer some novelty, surprise or incongruity in artworks, but not to much, there is a balance. He illustrates this with the inverted 'U' shape of the *Wundt curve*(Wundt, 1897), as depicted in fig. 2.12.

Berlyne's view takes into account the perceiver's past experience, and that as familiarity and expertise with a genre is gained, a work that previously was novel and interesting, can become boring. Similarly, a work that was previously perceived as too complex can become pleasurable with gained experience; a movement to the left along the Wundt curve.

2.6.3 Measures of Aesthetics Value

The *Golden Ratio* has been the subject of extensive interest, and has captured the public imagination. However its actual relationship to aesthetic value is debated. Markowsky lists a large numbers of misconceptions about the golden ratio, including debunking empirical results that claim the golden rectangle is the most preferred, that the human body exhibits these proportions, and that famous works of architecture such as the Parthenon and the great pyramids of Giza were built according to those proportions(Markowsky, 1992).

One of the most influential attempts at measuring aesthetic value is Birkhoff's *Aesthetic*

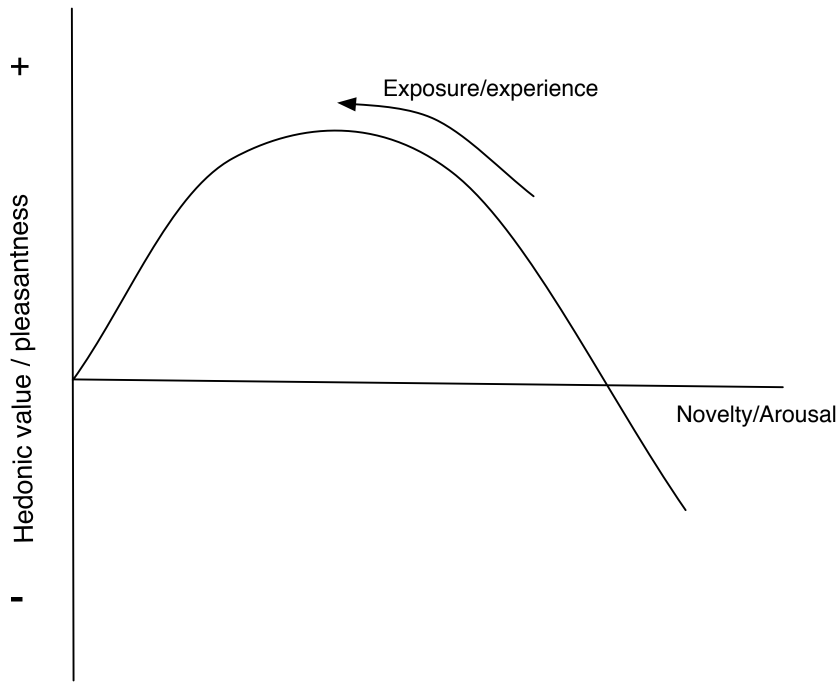


Figure 2.12: The ‘Wundt curve’. Objects of optimal aesthetic pleasure are at the top of the curve, with a moderate amount of novelty. Too little novelty can yield boredom. To much novelty can cause boredom or displeasure.

Measure. He defined the formula $M = O/C$, where M is the measure of aesthetic value, O is the order and C is the complexity. Birkhoff relates these terms to psychological subjective notions: the complexity C is the “preliminary effort of attention”, and O as “a realization that the object is characterized by a certain harmony, symmetry or order”, while the resulting M is “the feeling of value”(Birkhoff, 1933, p. 4).

How exactly to quantise this ‘order’ and ‘complexity’ of an object however is problematic and largely speculative. Subsequent researchers have studied the aesthetic measure, and have found often divergent and conflicting result(McWhinnie, 1968; R. Davis, 1936).

Despite this, the measure has been highly influential, not least on Bense with his ‘information aesthetics’, which as previously discussed, provided the context of some of the earliest computer art. Birkhoff’s measure is still explored to this day, and is being adapted and modified by researchers who seek to find a computable measure for aesthetic value (e.g (Rigau, Feixas, & Sbert, 2007; den Heijer & Eiben, 2010))

Another commonly explored principle is *Zipf’s law*, an observation that the frequencies of events in numerous naturally occurring phenomena are often distributed according to a power law relationship. First applied to natural languages, Zipf’s law states that the frequency of a word is inversely proportional to its rank in the frequency table. That is, the frequency of a given word P_i with rank i, is approximately equal to $1/1^a$, where a is close to 1(Newman, 2005). Manaris et al. have observed this distribution in the corpus of composers' works, and have developed systems to automatically classify works by a composer or style, and make predictions of 'pleasantness'(Manaris, Romero, et al., 2005).

Such objective measures (and others) of supposed aesthetic value have been applied by a number of researchers to define 'fitness functions' in evolutionary computation (e.g. (Manaris, Machado, McCauley, Romero, & Krehbiel, 2005; Reed, 2013; den Heijer & Eiben, 2010; Machado, Romero, & Manaris, 2008)).

2.6.4 Information theory aesthetics

It was noted in the survey of generative art and music, that Shannon and Weaver's information theory had a strong influence on the makers of computer art and music. Similarly, numerous researchers and theorists have applied it to the study of aesthetics in both art and music (e.g. (Youngblood, 1958; Coons & Kraehenbuehl, 1958; Moles, 1966; L. B. Meyer, 1967; J. E. Cohen, 1962)). The general view is that information theoretic quantities like entropy and redundancy are related to subjective sensations of uncertainty, surprise, complexity and tension.

Some theorists have combined other concepts such as Birkhoff's aesthetic measure with information theory, particularly as the later would appear to provide ways of quantising 'complexity' and 'order'(Rigau et al., 2007; Franke, 1971; Moles, 1966). There are numerous ways to quantify 'complexity' and 'order'. Informational complexity, as measured by Shannon entropy is one possible interpretation. Another is 'algorithmic complexity', which relates to the size of the smallest program that can reproduce the output, often known as Kolmogorov complexity(Li & Vitányi, 2013).

Another was proposed by Gell-Mann, *effective complexity*(Gell-Mann & Lloyd, 1996). It draws from the observation that measures such as entropy do not coincide well with human perception of complexity. White noise does not sound as complex to human ears as a Bach fugue, yet in terms of entropy, white noise is maximal. Similarly the most complex entities, living things, do not have maximum disorder, but reside somewhere in-between complete order (zero entropy) and complete disorder, at the so-called "edge of chaos"(Langton, 1990). As such effective complexity "corresponds to our intuitive notion of complexity"(Gell-Mann & Lloyd, 1996).

In some ways effective complexity is reminiscent of Berlyne's theory. Correlating order/disorder to arousal, both hedonic response and effective complexity are minimal at the extremes and maximal at intermediate amounts. Drawing on this observation, Galanter advocates for effective complexity "as a guiding principle in aesthetics"(Galanter, 2003).

Information Dynamics of Music

Information theoretic approaches to understanding the musical experience have also been explored. Meyer (L. B. Meyer, 1967) and Narmour (Narmour, 1977) argued that the musical experience is highly dependent on how listeners continually build up expectations on future events, and how these expectations are fulfilled or deceived are key to the perceived aesthetic value of the work.

There is evidence that listeners internalise statistical knowledge about musical structure(Saffran, Johnson, Aslin, & Newport, 1999), and also that statistical models can form an effective basis for analysis of music (Conklin & Witten, 1995; Ponsford, Wiggins, & Mellish, 1999; Pearce, 2005).

Building on this, *information dynamics*(Pearce & Wiggins, 2012; Abdallah & Plumbley, 2009; Potter, Wiggins, & Pearce, 2007) builds a cognitive account of expectation and surprise in music, and studies these using the tools of information theory. Rather than considering music as a static object presented in its entirety, information dynamics focuses on subjective expectation and surprise *as they happen*.

Information dynamics views listener expectations as being based on both the familiarity with musical styles, but also on the listener's ability to detect and learn statistical regularities in the music as they emerge; listeners continually revise this internal belief state as the music unfolds, which includes predictive distributions over possible future events. These changes in probabilistic beliefs can be associated with quantities of information; the focus of information dynamics.

Information dynamics informed the design of the first practical exploration of this thesis, the *Melody Triangle*, presented in the next chapter.

2.6.5 Computational Aesthetics

New media inevitably alter the shape of art; they bring about new possibilities and extend the range of expression for the creative people that employ them. Accordingly the introduction of computers and digital media has changed what artists and musicians do. Computers however are not only used to shape or construct artefacts, they can also be used to shape or construct *ideas*.

Because computers required that ideas are specified with precision, computers help us to “think as well as make”(Candy & Edmonds, 2002, p. 28). Hence, they can be expected to change the way artists think about their work as well as changing the form of the work. As Candy and Edmunds suggest, “we see the impact of the digital world on the artist falls into two areas: on thinking about art (development) and making art (delivery)”(Candy & Edmonds, 2002, p. 28).

As the many examples mentioned in this chapter suggest, generative processes and computation offer new possibilities for the realisation forms and patterns. But as the artist is confronted with technological conditions, it can engender new thinking and raise questions about the nature of art and the creative act itself. Artistic acts are traditionally biased on past experience; not just in the pre-conceptions of form and structure, as embodied in cultural styles and conventions, but these pre-conceptions also define the roles of the artist, the artwork and the audience in a social context. Perhaps the true gift of generative art is it introduces a mathematical stem that is disconnected from these kinds of conceptions; and with it brings the opportunity for a computational aesthetics that challenges these patterns of thinking and traditional assumptions.

Joseph Schillinger, in his book *The Mathematical Basis of the Arts* expresses one such challenge to traditional assumptions:

> “If art implies selectivity, skill and organisation, ascertainable principles must underlie it. Once such principles are discovered and formulated, works of art may be produced by scientific synthesis. There is a common misunderstanding about the freedom of an artist as it relates to self-expression. No artist is really free. He is subjected to the influences of his immediate surroundings in the manner of execution, and confined to the material media at hand. If an artist were truly free he would speak his own individual language. In reality he only speaks the language of his immediate geographical and historical boundaries ... The key to real freedom and emancipation from local dependence is through scientific method ... Creation directly from *principles*, and not through the *imitation* of appearances is the real way to freedom for an artist”(Schillinger, 1948, p. 3).

This perhaps a fringe and extreme view, and a decisively un-romantic one. Where is the room for emotion, intuition and imagination? Or is he correct in supposing that the ultimate destiny of art is its production by ‘scientific synthesis’?

A fringe view perhaps, but as some of the research surveyed in this chapter suggests, it is not so uncommon to find researchers that, through experimentation and careful measurement and analysis of creative artefacts, seek to find amongst the bewildering complexity of the world of art and music what factors, what parameters are important, and with them attempt to create an automation of art and music. In fact, the practical elements of this research possess a similar alignment; they also seek to find patterns in parameters and uncover their relationship to aesthetic judgments.

The idea of computers creating art or music has been known to be met by resistance, not least as it is at odds with the traditional romantic view that regards music and art as direct communication of emotion from artist to contemplator. Well known cases of this include the reactions to Cohen's drawing and colour program *AARON*(H. Cohen, 1995). *AARON* has created artefacts and exhibited at venues all around the world including the Tate Gallery. An almost equally famous example is *Emmy*, a composition system developed by the composer David Cope(Cope, 2004, 2005). *AARON* and *Emmy* represent an extreme cases of generative art; the meta-creation of art through the creation of an anthropomorphised generative system. However many of the philosophical issues they raise are applicable to generative art more broadly.

Can a computer really be truly autonomous and responsible for creating a work of art/music? Can these autonomous creations truly be creative? Or is the programmer, however indirectly the true creator of the work?

In one respect it may seems that this is a problem that is dependent on *where* the generative process lies. If a generative process is designed and carried out entirely by a human actor, as some of the supposed earliest generative art was(Galanter, 2003), then it seems fair to attribute authorship to the actor. But sometimes the generative rules – or constraints – are not the defined by the same actor as the one who carries them out. Considering for instance Schoenberg's rules of serialism; should all serialist compositions be, at least in-part, attributable to Schoenberg? What if the rules and constraints are specified and embodied in a computer as is the case for *AARON* and *Emmy*? And if they make something that their creators did not anticipate, could they be called 'creative'? And what if Cope and Cohen had not chosen to anthropomorphise their creations and give them names? These are but some of the many questions of creative responsibility and attribution that get raised in computational aesthetics, however the issues do not end here.

Traditionally in the world of fine-art, one of the issues concerning aesthetic evaluation relates to *authenticity*. Masterfully crafted fraudulent copies of, say, the *Mona Lisa* are not considered of value, even if for all intensive purposes they look identical. If a computer can generate hundreds or thousands of instances of a work at the press of a button, can they maintain an ‘authenticity’? Cope has suggested that one of the reasons the works generated by *Emmy* were ill received had to do with the fact that at the push of a button *Emmy* could be made to generate endless more such works(Cope, 2005). That flesh and bone composers are mortal, and their oeuvre is as such bounded, has consequences for the aesthetic evaluation.

This ‘immortality’ of generative works however is an illusion. Even though some of the earliest computer artists such Georg Neese and Frieder Nake could – at the time – command their computers and printers to create endless instances of any of their works, this no longer possible: the systems that generated the earliest computer art no longer exist. Computer generated artworks may seem to have a capacity for endless re-generation, but they do not in practice, because they are tied to a specific technological substrate. All technology goes obsolete, the older a piece of software or hardware is, the harder it is to find the requisite parts and compatible technology to run it today. In recent years the conservation of digital artworks has become recognised as a concern(Serexhe, 2013). Does a work become more ‘authentic’, and hence valuable, if it is difficult or impossible to reproduce?

These philosophical waters are muddied even further when considering works that are open to interference due to input from the environment or the ‘consumers’ of the work. This is the case with interactive works, which all the practical studies of this thesis are. If an audience member’s action are responsible for the outcome of the form of the work, where does the creative responsibility lie? And is the audience truly still just an audience if they have such an impact on the work? These questions are very relevant to the three practical explorations of this thesis, as all take in input from an ‘audience’ to shape the direction and outcomes of the work.

For evolutionary based art works, such as the many works of a-life, if an artists specifics the ‘fitness function’ that the work optimises towards, then perhaps the attribution of the work lies with the artist. However if the fitness function is defined through some sort of interaction with an audience, as is the case with the final practical study of this research – *EvoColour* – where can the credit for the generated works lie? The generated forms are after all the consequences of both the programmer and the audience.

It may be the case that this ambiguity of authorship is precisely the point. Roy Ascott for instance suggests that the value in interactivity is as a democratisation of the creative process - "creativity is shared, authorship is distributed"(Ascott, 1990). If authorship is deliberately distributed, then one need not worry about ascribing it to any one source.

Besides what is the 'artwork' in this scenario anyway? Is it the software that generated the work? Or is it the generated artefacts themselves, be it prints, video or even music, that is the 'artwork'? Or is it the generative algorithm as some abstract entity?

In short, computational aesthetics poses many difficult questions of authenticity, agency, creative responsibility and ontology. Boden and Edmunds propose a way of side-stepping many of these sticky philosophical issues: "perhaps we should speak not of the 'artwork' but of the 'art system' – where this comprises the artist, the program, the technological installation (and its observable results), and the behaviour of the human audience? (And perhaps, if the concept of the 'artwork' falls, then that of the 'artist/author' falls too?)"(Boden & Edmonds, 2009)

Indeed it may well be that the traditional view of art having a clearly defined author/composer, that is the source of the creativity and authenticity of the work, and that realises the work in a gallery or concert hall for consumption by a clearly demarcated audience, is obsolete. And indeed, perhaps the practical works of this thesis are perhaps best considered to be 'art systems'.

This thesis carves out a small territory in this larger field of computational aesthetics. The focus is on parameter search, and on the dynamics of aesthetic value; as parameter values are measurable, they can provide the means to probe analytically the nature of aesthetic value judgments. The detected patterns of parameters can have practical applications, perhaps even take steps towards a 'scientific synthesis' of art. But rather than just us the computer for the industrialisation of art and music, as Schillinger would have us do, perhaps, as John Cage suggests "what we need is a computer that isn't labour-saving, but which increases the work for us to do."(Cage, 1970).

Chapter 3

The Melody Triangle

3.1 Introduction

The *Melody Triangle* is the first practical exploration of this research. It is an interface for the discovery of musical content where the parameter space of a stochastic generative musical process, the Markov chain, is 'mapped out' according to the *predictability* of the generated output. The *Melody Triangle* was developed in the context of *information dynamics*(Pearce & Wiggins, 2012; Abdallah & Plumbley, 2009), already discussed and contextualised in section 2.6.4 of the background chapter.

The *Melody Triangle* interface provides a mapping between the *quality dimensions*(Gärdenfors, 2004b) of subjective predictability, and the input parameters to Markov processes. This mapping allows users of the triangle to explore Markov processes *in terms* of the predictability quality dimensions. For instance, one corner of the triangle returns completely random and unpredictable melodies, while an other area yields predictable and periodic patterns; the entirety of the triangle covering a spectrum of predictability.

Like the other practical explorations of this thesis, parameter values selected by users interacting with the *Melody Triangle* are collected, and then analysed with a view of finding patterns in aesthetic preferences. Observing what areas of the triangle are most popular with users could reveal how the predictability quality dimensions relate to user preferences; in other words - how predictable do we like our music?

There are many potential applications and benefits to building such a map relating predictability to aesthetic value. Such a map could inform the design of algorithmic composition systems for instance, which could be made to ensure that the generated content is of just the right amount of 'novelty'; not too boring because of excessive repetition, and similarly not too boring due to a lack of coherence and structure. Further such a map could provide a framework for the automatic evaluation of musical works, and help predict how well they would be received by a human listener.

In section 3.2 the information measures that lead to the development of the *Melody Triangle* are outlined. In section 3.3 how these information measures are used to construct the *Melody Triangle*, and how the triangular interface is used to retrieve patterns of symbols (which are then mapped to notes or percussive sounds) is described.

The *Melody Triangle* has been implemented as an interactive installation, a desktop application and as a mobile phone application for the Android operating system, each new incarnation building on lessons learned from the last. The interactive installation, described in 3.4 was the first incarnation of the *Melody Triangle*. It was premiered at Digital Shoreditch[1], and has been exhibited at the Brighton Science Festival[2] as well as the Bradford Science Festival[3]. In the installation, multiple visitors would generate music by exploring a triangular area in the room. When in the triangle a melody would be generated by the system and played back in the space. As with all incarnations of the *Melody Triangle*, the position in the triangle determined the predictability of the output melody. A Kinect camera was used to track the individual visitors, and additionally could detect gestures that allowed the visitors to change the register, instrumentation and rhythmic attributes of their melody.

A study with the installation to determine what areas of the triangle are most popular was considered, however it was clear from the behaviour of visitors that any data collected would be too noisy to derive significant observations; the situation was too chaotic and playful. This then motivated the implementation of the desktop incarnation of the *Melody Triangle*, described in 3.5. Here a user would control the music using a mouse and keyboard, dragging tokens representing melodies on screen in and around the triangle, with some key commands to change other musical parameters. Qualitative feedback from practicing musicians (summarised in 3.5.4) indicates that the interface has potential as a useful composition aid or live performance tool.

[1]http://digitalshoreditch.com/
[2]http://www.brightonscience.com/
[3]http://www.bradfordsciencefestival.co.uk/

The main purpose of the desktop interface was for use in a study described in 3.5.1. In this study participants would explore the musical properties of the triangle by repeatedly picking a single position in the triangle, and indicating wether they 'liked' what they heard by pressing the spacebar. The data collected did not contain statistically significant patterns, due to a number of flaws in experimental design. However feedback from the subjects did indicate that they were able to identify the qualities of the predictability dimension as mapped to the different areas of the triangle.

This in turn motivated the implementation of the *Melody Triangle* mobile phone application. In section 3.6, its features are outlined, including how it allows users to share their settings with each other, and how it has been used to collect data. As users of the app find melodies and patterns they like, they are encouraged to press a 'like' button, where their setting are uploaded to a server for analysis. The collected 'liked' settings of users worldwide were used to find trends and commonalities in across the submissions of the users of the app, and to investigate how these might relate to the aesthetic preferences and the information-dynamic model of human expectation and surprise.

As will be shown in the section 3.6.3, users of the mobile app clearly preferred the more predictable side of the predictability quality dimension. However there were some indications that the design of the interface carried within it some affordances that muddied the results; a strong tendency to select places in the triangle that corresponded to visual symmetry made it unclear if those areas were selected for their musical attributes or their visual properties. A number of challenges and lessons to be drawn from this are articulated, and proposals for facilitating such data collection is proposed.

In this next section how the *Melody Triangle* is constructed and the information measures that define it are described.

3.2 Information Measures

As outlined in the discussion in section 2.6.4 of the background chapter, information dynamics approaches musical experience by modelling a listener experiencing music *in time*. With this view, a listener continually makes predictions about how the piece will continue, based on previous exposure[4]. The working hypothesis is that as a piece of music is heard, the listener revises

[4]Which are both cultural, such as with regards to exposures to musical styles, as well as with respect to the immediate context of the musical piece at hand

an internal state, which includes anticipations and predictions over possible future events(Pearce & Wiggins, 2012; Abdallah & Plumbley, 2009).

Shannon and Weaver(Shannon & Weaver, 1964) used the metaphor of information transmitted from a 'source' over a communication 'channel' to a 'receiver'. An information dynamic approach however re-casts this channel as a time-channel, not a physical transmission channel. As the receiver/listener applies some algorithms to predict the current sample from its past, Dubnov explains, "the transmission process over a noisy channel has now the interpretation of prediction/anticipation performance over time"(Dubnov, 2006).

Some fundamental information measures are first very briefly reviewed - these are *conditional entropy* and *mutual information*. These are then used to define the measures that defined the triangle - *entropy rate* and *redundancy*. Additionally we also describe the measure of *predictive information rate*(Abdallah & Plumbley, 2009), that although not strictly necessary in the construction of the triangle, provides valuable insights into the phenomenological effects of the system's generated outputs.

3.2.1 Conditional Entropy, Mutual Information

The *entropy* of a random variable X, written as $H(X)$, represents the observer's uncertainty about X. *Conditional entropy* is defined as the entropy of a random variable Y *conditioned* on another X, written as $H(Y|X)$, and depicts the uncertainty about X if Y is known.

Mutual information between two variables X and Y, $I(X,Y)$ is defined as:

$$I(X,Y) = H(X) - H(X|Y) \tag{3.1}$$

$I(X,Y)$ represents the amount of *shared information* between the two variables. That is, how much information one variable carries about the other. This represents both the expected information in an observation of Y about X and the expected reduction in uncertainty about Y after observing X.

Further details on these basic measures of information can be had from any basic textbook on information theory. In the next sections the measures relevant to the *Melody Triangle* - *entropy rate*, *redundancy* and *predictive information rate* - are discussed. More detailed definitions and derivations of these measures can be found in (Abdallah & Plumbley, 2009, 2010; Dubnov, 2006).

3.2.2 Entropy Rate

A sequence of symbols from the viewpoint of an observer at a certain time, t, can be split into a single symbol in the *present* X_t, an infinite *past* $\overleftarrow{X}_t \Leftrightarrow (\ldots, X_{t-3}, X_{t-2}, X_{t-1})$, and an infinite *future* $\overrightarrow{X}_t \Leftrightarrow (X_{t+1}, X_{t+2}, X_{t+3}, \ldots)$. The symbols are arriving at a constant uniform rate.

The *entropy rate* is a well-known measure of randomness or unpredictability of a sequence, and is directly related to the measure of conditional entropy.

The entropy rate is the conditional entropy, H, of the *present* given the *past*:

$$h_\mu = H(X_t|\overleftarrow{X}_t) \tag{3.2}$$

Entropy rate represents the averaged uncertainty about the present symbol *given* that we have observed everything before it. In other words, it is the *time average* of the conditional entropies. Processes with zero entropy rate can be predicted perfectly given enough of the preceding context. μ is used here to represent this process of calculating the long-run time average.

3.2.3 Redundancy

The mutual information I, between the 'past' and the 'present' was defined by Dubnov(Dubnov, 2006), as *information rate*:

$$\rho_\mu = I(\overleftarrow{X}_t; X_t) = H(X_t) - h_\mu = H(X_t) - H(X_t|\overleftarrow{X}_t). \tag{3.3}$$

Abdallah and Plumbley refer to this measure as the *multi-information rate*(Abdallah & Plumbley, 2010) ρ_μ - is a time average of the mutual information, and can be thought of as a measure of *redundancy*: it quantifies the extent to which the same information is to be found in all parts of the sequence. *Redundancy* is the more descriptive term and as such will be used here on in to refer to ρ_μ. Redundancy is the difference between the entropy of a single element of the sequence in isolation, and its entropy after taking into account the preceding context; if the previous symbols reduce our uncertainty about the present symbol a great deal, then the redundancy is high.

For example, for a sequence that consists of a continuously repeating cycle such as ...b, c, d, a, b, c, d, a ..., then the redundancy ρ_μ is high. This is because as $H(X_t)$ is high, when considering symbols in isolation without context, while $H(X_t|\overleftarrow{X}_t)$ is zero, because knowing the previous symbol immediately tells us what the present symbol is (once the listener has learned the sequence). In other words, this long term average of mutual information - *redundancy* - "depends on the surprise that the time-channel introduces to the next sample versus the ability of the listener to predict this surprise."(Dubnov, 2006)

3.2.4 Predictive Information Rate

Abdallah and Plumbley in (Abdallah & Plumbley, 2009) define the measure of *predictive information rate PIR*. PIR brings the uncertainty about the future - $\overrightarrow{X}_t$ - into play. Although it is not explicitly needed in the construction of the *Melody Triangle*, it is outlined here as it gives valuable insight into the perceptual properties of the predictability quality dimensions that the triangle helps users explore. PIR is a measure of how much each symbol reduces the uncertainty about the future as it is observed, *given* that we have observed the past:

$$b_\mu = I(X_t; \overrightarrow{X}_t | \overleftarrow{X}_t) = H(\overrightarrow{X}_t | \overleftarrow{X}_t) - H(\overrightarrow{X}_t | X_t, \overleftarrow{X}_t). \tag{3.4}$$

It is a measure of the mutual information between the 'present' and the 'future' given the 'past'. In other words, it is a measure of the average amount of *new* information in each symbol. As such the PIR is low for regular, repetitive sequences, as there is no new information in each subsequent observation. Similarly the PIR is also low for random sequences, as again each observation carries no new information about the sequences to come. Instead processes with high PIR maintain a certain kind of balance between predictability and unpredictability.

3.2.5 Information Properties of Markov Processes

In experiments visualising and sonifying sequences sampled from first order Markov chains, Abdallah and Plumbley found that the measures of entropy rate (h_μ), redundancy (ρ_μ) and PIR (b_μ) correspond to readily perceptible characteristics; a high entropy rate is associated with completely uncorrelated sequences with no recognisable temporal structure (and low ρ_μ and b_μ), and that high values of redundancy are associated with long periodic cycles (and low h_μ and b_μ). And further that high values of PIR correlate with intermediate values of entropy rate and redundancy, and yield recognisable, but not completely predictable, temporal structures(Abdallah & Plumbley, 2010).

Fig. 3.2 is a 3-D scatter plot of the information measures for hundreds of randomly sampled[5]

[5]The random sampling was weighted to increase the probability of generating sparse transition matrices, to get a good spread that reaches the edges and corners

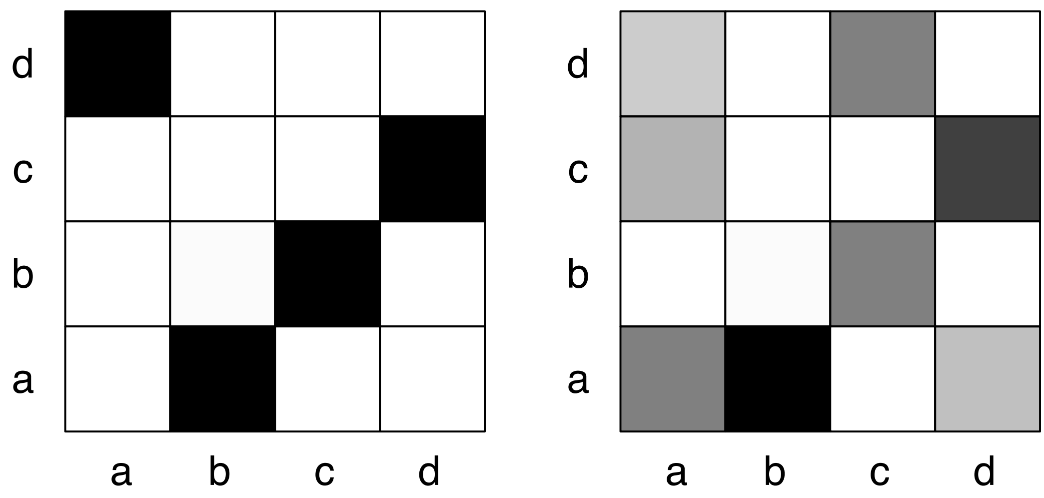
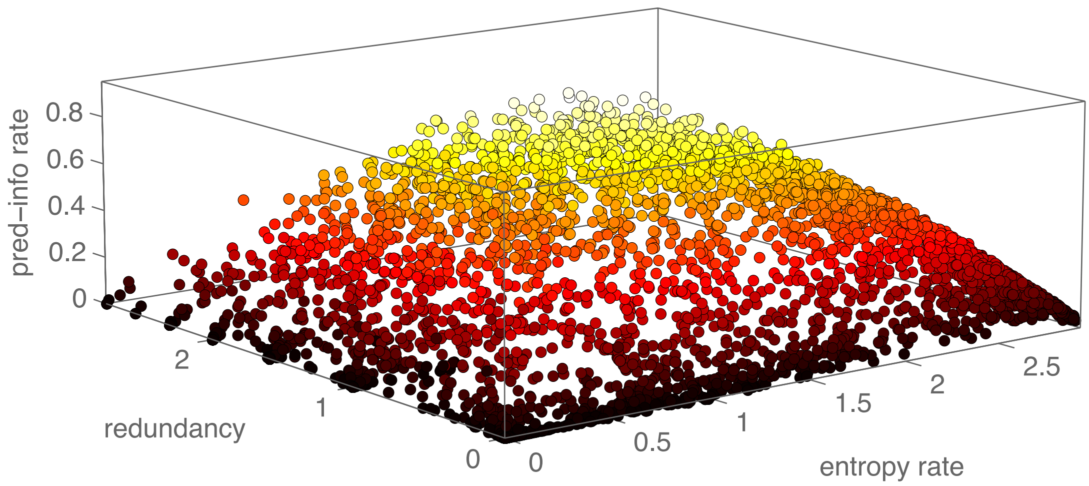
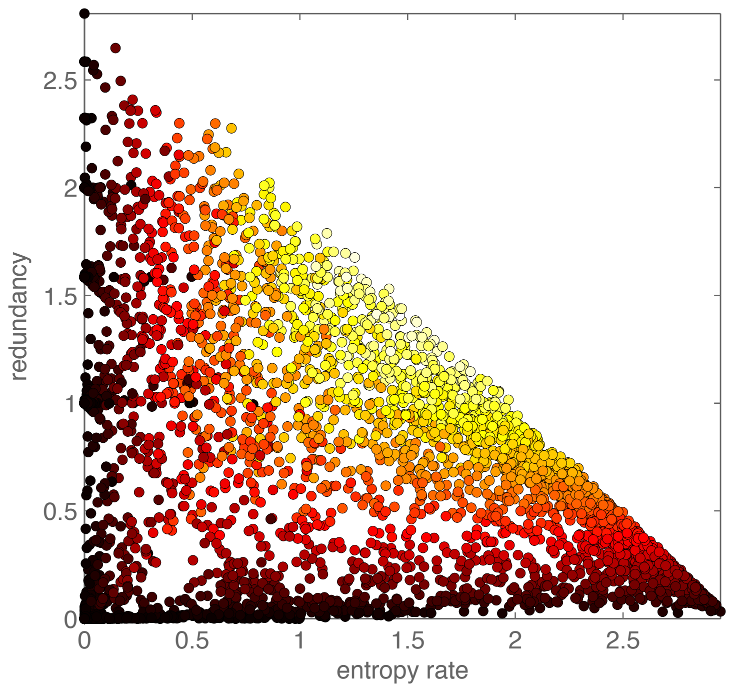


Figure 3.1: Two transition matrices representing Markov chains. The shade of grey represents the probabilities of transition from one symbol to the next (white=0, black=1). The current symbol is along the bottom, and the next symbol is along the left. The left hand matrix has no uncertainty; it represents a periodic pattern (...,a,d,c,b,a,d,c,b,a,d,c,b,a,...). As such it has high redundancy, but low entropy rate, and low predictive information rate. The right hand matrix contains unpredictability but nonetheless is not completely without perceivable structure (we know for instance that any 'b' will always be followed by an 'a' and preceded by a 'c'), it is of a higher entropy rate and predictive information rate, but lower redundancy.



(a)



(b)

Figure 3.2: Distribution of hundreds of 8-state transition matrices in the 3D space of entropy rate (h_μ), redundancy (ρ_μ) and predictive information rate (b_μ), all in bits. As can be seen in (a) the distribution as a whole makes a curved sheet, with the highest PIR values found at intermediate entropy and redundancy. Although not visible in this plot, it is largely hollow in the middle. As can be seen in (b), the same plot with the PIR dimension projected out forms a right angled triangle, this is the triangle which corresponds to the interface of the *Melody Triangle*. The concentrations of points along the redundancy axis correspond to Markov chains which are roughly periodic, with increasing periods the higher the redundancy; in this case period 7 (redundancy 2.8 bits)

eight state Markov chain transition matrices. The coordinates of the 'information space' are entropy rate (h_μ), redundancy (ρ_μ), and predictive information rate (b_μ). The points along the redundancy axis correspond to periodic Markov chains, where as those along the entropy axis produce uncorrelated sequences with no temporal structure. As can be seen, processes with high PIR are to be found at intermediate levels of entropy and redundancy. PIR is low both for regular processes, such as constant periodic sequences, *and* low for random processes, where each symbol is chosen independently of the others.

Abdallah and Plumbley note how this balance between predictability and unpredictability of the PIR is reminiscent of the inverted 'U' shape of the Wundt curve(Abdallah & Plumbley, 2009). Already entioned in section 2.6.4 of the background chapter, the Wundt curve suggests that stimuli are most pleasing at intermediate levels of novelty or disorder, where there is a balance between 'order' and 'chaos'.

An inverted 'U' shape is visible in the upper envelope of the plot in fig. 3.2a. The distribution of the transition matrices in this space form a thin curved sheet, and is hollow inside. This comparison with the Wundt curve a suggestion that the PIR might be related to the Berlyne's 'hedonic value'(Berlyne, 1970).

It was observed how the natural distribution of Markov processes plotted along entropy rate and redundancy formed a triangular shape as shown in fig. 3.2b. It is this triangular shape that then formed the interface that became the *Melody Triangle*, this is explained in greater detail in 3.3.

3.3 Constructing The Melody Triangle

The *Melody Triangle* is musical interface that is designed around the natural distribution of Markov chain transition matrices in the information space of entropy rate (h_μ) and redundancy (ρ_μ), the right-angled triangle, as illustrated in Fig. 3.2b.

The right-angled triangle is rotated and stretched to form an equilateral triangle with the 'redundancy'/'entropy rate' vertex at the top, the 'redundancy' axis down the right-hand side, and the 'entropy rate' axis down the left, as shown in fig. 3.3. This is the *Melody Triangle* and forms the interface by which the system is controlled.

The corners correspond to three different extremes of predictability and unpredictability, which could be loosely characterised as 'periodicity', 'noise' and 'repetition'. Melodies from

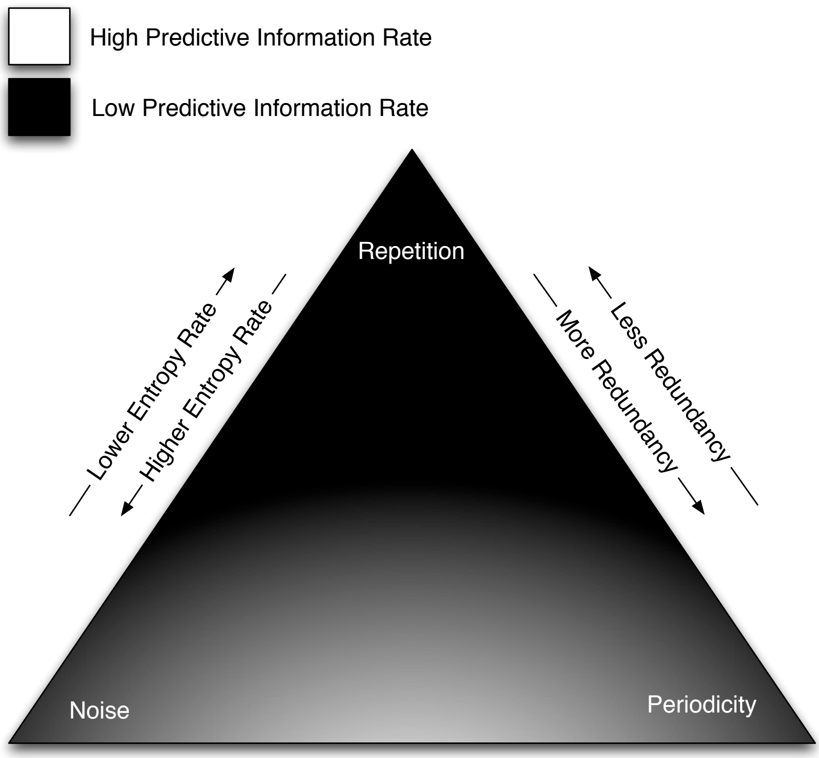


Figure 3.3: The *Melody Triangle* with its relationship to entropy rate (maximum at the bottom left), redundancy (maximum at the bottom right) and predictive information rate (maximum in middle bottom). The ‘repetition’ corner maps to sequences of a single repeated note. The ‘periodicity’ corner maps to long loops, and ‘noise’ corner maps to completely random sequences.

the ‘noise’ corner (high h_μ, low ρ_μ and low b_μ) have no discernible pattern; those along the ‘periodicity’ to ‘repetition’ edge are all cyclic patterns that get shorter as one approaches the ‘repetition’ corner, until each is just one repeating note. Those along the opposite edge consist of independent random notes from non-uniform distributions. Areas between the left and right edges will tend to have higher predictive information rate (b_μ), and it is hypothesised that, these will be perceived as more ‘interesting’ or ‘melodic’. These melodies have some level of unpredictability, but are not completely random. Or, conversely, are predictable, but not entirely so.

3.3.1 The Melody Triangle as a Conceptual Space of Predictability

In arranging Markov chains according to information measures in this way, the *Melody Triangle* provides a mapping between the parameters of the Markov processes, and how predictable it will be perceived to be. Conversely if one were to plot the Markov chains according to their parameter values, as defined by the values in their transition matrices, some chains that are close to each other in parametric space would not have similar perceptual properties, while some far apart in parameter-space would have very similar properties. There is nothing about the euclidean distance between those parameter values that would correspond to similarity.

Gärdenfors' theory of *conceptual spaces*(Gärdenfors, 2004b), outlined in section 2.4.2 of the background, is a framework for the representation of concepts, and supports spatial reasoning in terms of similarity; objects that are phenomenally similar are close to each other in the conceptual space defined by a number of *quality dimensions*.

The *Melody Triangle*, with its axes of entropy rate, and redundancy displays many of the characteristics of a conceptual space. Markov chains spatially close to each other in this space have similar phenomenal characteristics, and certain geometries of the triangle could be said to correspond to 'natural concepts'. The right edge, for instance, represented the natural concept of the 'perfect loop', the top corner represents the natural concept of 'repetition', the left edge corresponds to natural concept of 'perfect randomness'. The left-right direction could be said to correspond to the 'quality dimension' of predictability, and the up-down axis corresponds to the quality-dimension of 'many/fewness', as the further down one goes, the more symbols/notes are involved in the pattern on average.

There are however some significant caveats. The information measures that define the *Melody Triangle* assume a constant rate of symbols, and thus the output sequences proceed at a constant, uniform rate. Although the placing of events in time and rhythm has a strong effect on expectations, surprise and satisfaction in music, the system does not address this temporal dimension. Additionally the system does not address the culturally defined expectations of melodic structure that result from our exposure to tonal music; all symbols are considered equal, regardless of what note in a scale they are mapped to.

Nevertheless it could be said that the *Melody Triangle* represents a 'conceptual space of predictability' *for first-order Markov processes*, and as such, it affords reasoning about the predictability of these processes in spatial and geometric terms.

3.3.2 Usage and Sonification

The different incarnations of the *Melody Triangle* – the interactive installation, the desktop application and the mobile app – all work in fundamentally the same fashion. Thousands of Markov chains are generated, and then mapped according to their entropy rate and redundancy values[6].

[6] Although the triangle is 2d, the third dimension, the PIR (b_μ), is present implicitly, as the transition matrices along the centre line will tend to higher PIR.

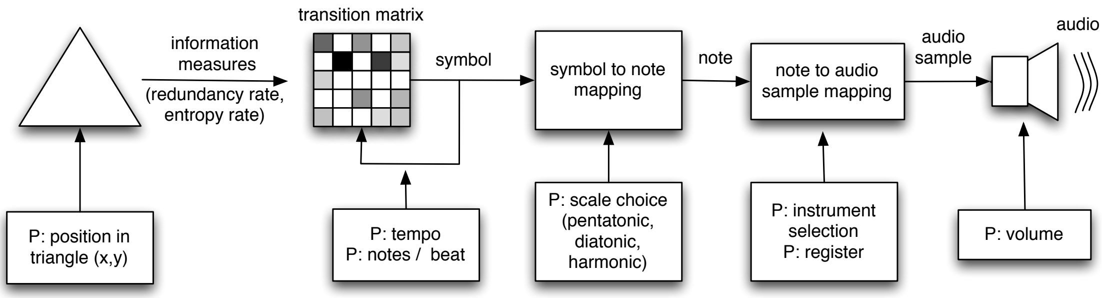


Figure 3.4: Input parameters and how they lead to the generation of audio. Parameters are indicated by *P*. The position in the triangle map to information measures. These in turn cause a Markov transition matrix with these information properties to be selected. A sequence of symbols is then generated based on the matrix, at a rate based on the tempo and notes per beat parameters. These are mapped to notes based on the scale parameter. The notes are mapped to audio samples based on instrument selection and register parameters.

A user selects a point within the triangle, and this position maps into the information space of entropy rate and redundancy. How this selection is made depends on the incarnation: for the installation this is done by placing one's body in the space, for the desktop application and mobile app this is done by dragging a round token into the triangle through a graphical interface.

The nearest Markov chain in this information space is used to generate a sequence of symbols which are then sonified either as pitched notes or percussive sounds. This process is outlined in figure 3.4.

In each of the three incarnations of the Melody Triangle, the musical sounds are encoded and generated in the same way. The *Melody Triangle* has a core tempo (which can be adjusted by the user). When a Markov chain has been selected by specifying its position in the triangle, at each time-step, a symbol is generated. The sequence of symbols correspond to the chain being 'played out'. Each generated symbol is fed-back as input to the transition matrix, which defines the probability of the subsequent symbol, which in turn is the input to the matrix for the subsequent symbol, and so on. The subjective predictability of the generated sequence is thus determined by the position in the triangle. The symbols are generated at a rate that is an integer multiple of the core tempo; a 'notes per beat' value associated to each token, and controllable by the user.

The audio is generated by then mapping each of these symbols to an audio sample. In the installation and the desktop application, this is implemented by sending MIDI messages that trigger audio in an external audio application (Apple's *Logic*). Whereas in the mobile app, the

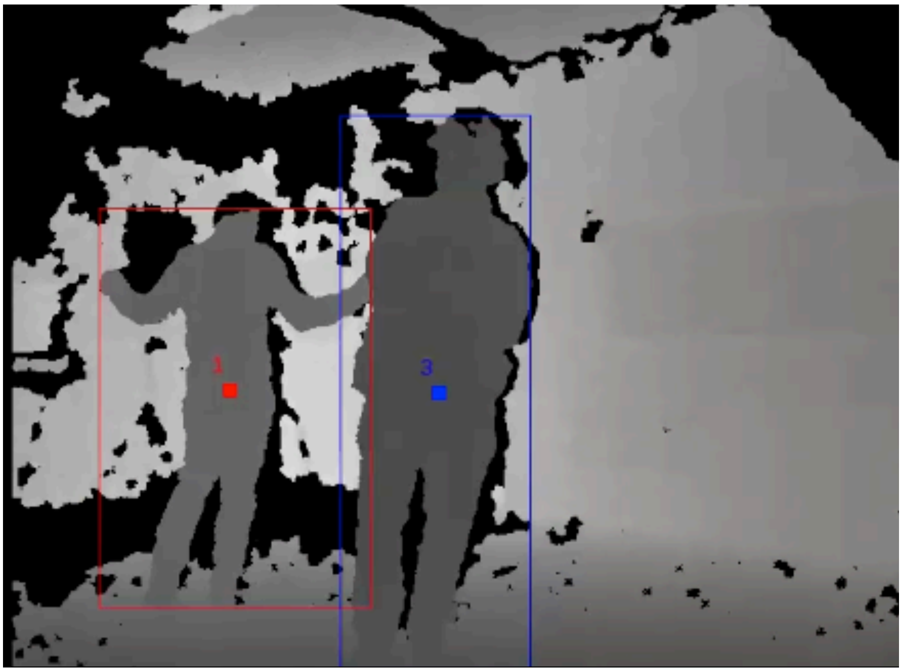


Figure 3.5: The depth map as seen by the Kinect camera, and the bounding box outlines the blobs detected by OpenNI.

audio sample playback was in-built into the app[7].

The user can choose between banks of audio-samples that the symbols map to. In the three incarnations, the banks of audio samples corresponded to three instruments; percussion/drums, piano and bass. For the pitched audio samples (piano and bass), the user of the triangle could also specify a musical scale (such as pentatonic, diatonic, harmonic or chromatic), which would thus determine which symbol was mapped to which audio sample. Additionally, the user can specify an octave range for the pitched samples, allowing the instruments to play in different registers.

In the next section the first incarnation of the *Melody Triangle*, the interactive installation, is described.

3.4 The Interactive Installation

A Kinect[8] camera was used to track individual visitors in the installation. Kinect cameras are able to detect how far away objects are by building a 'depth map' (see fig. 3.5), and doing so, its range naturally forms a triangle. As visitors/users comes into the range of the camera, they

[7]Although in principle those symbols need not be mapped to audio at all, and could be used to trigger sequences of colours for instance, a process that occurs in the generation of images for *EvoColour* described in chapter 5.

[8]A motion sensing input device developed for the Microsoft Xbox game consoles, commonly used in interactive art contexts.

Table 3.1: Gestures and their resulting effect in the Melody Triangle installation

left arm	right arm	meaning
out	static	double tempo
in	static	halve tempo
static	out	triple tempo
static	in	one-third tempo
out	in	shift to off-beat
out	out	change instrument
in	in	reset tempo

start generating a melody, the statistical properties of this melody determined by the mapping of physical space to statistical space as discussed above. Thus by exploring the physical space the participant changes the predictability of the generated melodic content. When multiple people are in the space they can cooperate to create interweaving melodies, forming intricate polyphonic textures.

The streams of symbols are mapped to MIDI and then played with software instruments. The tracking system was capable of detecting gestures, and these were mapped to different musical effects such as tempo changes, periodicity changes (going to the off-beat), instrument/register changes and volume (see table 3.1, figure 3.6).

3.4.1 Tracking and Control

Tracking and control was done using the OpenNI libraries' API[9] and high level middle-ware for tracking with Kinect. This provided reliable blob tracking of humanoid forms in 2d space. By triangulating this to the Kinect's depth map it became possible to get reliable coordinate of visitors' positions in the space.

By detecting the bounding box of the shape of the individuals in the space, and then normalising these based on the distance of the depth map, it became possible to work out if an individual had an arm stretched out or if they were crouching. A series of gestures for controlling the system without the use of any controllers was then defined(see table 3.1). For instance, by sticking out one's left arm quickly, the melody doubles in tempo. By pulling one's left arm in at the same time as sticking the right arm out the melody would shift onto the offbeat. Sending out both arms would change the instrument being 'played', crouching would decrease the volume, while standing tall and raising up the arms would increase it.

[9] http://OpenNi.org/

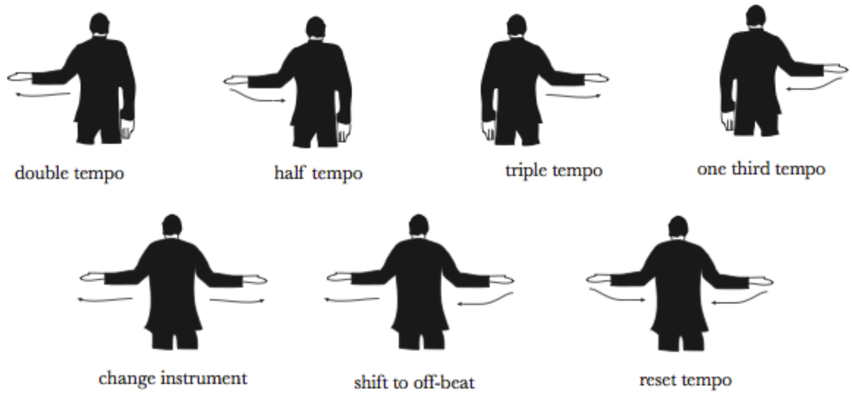


Figure 3.6: Gestures and their resulting effect

3.4.2 Observations

Although visitors would need an initial bit of training they would then quickly be able to collaboratively design musical textures. For example, one person could lay down a predictable repeating bass line by keeping themselves to the periodicity/repetition side of the room, while a companion can generate a freer melodic line by being nearer the 'noise' part of the space. By not having one user be able to control the whole narrative, the participants would communicate verbally and direct each other in the goals of learning to use the system and finding interesting musical textures. This collaboration added an element of playfulness and enjoyment that was clearly apparent.

This installation is an exploratory prototype and occupies an ambiguous role in terms of purpose; it is in a nebulous middle ground between instrument, art installation and technical demonstration. It is clear however, that as a vehicle for communicating ideas related to the expectation, pattern and predictability in music to the public, it proved very effective[10].

It was hoped that the installation could be used to identify what areas of the triangle were most popular. However it was clear from observing how people behaved in the installation that this would yield very noisy data unlikely to be of much use; the situation was too playful and chaotic. In order to better get a sense of preferences with regards to the information models, the desktop version of the *Melody Triangle* was developed. This is described in the next section.

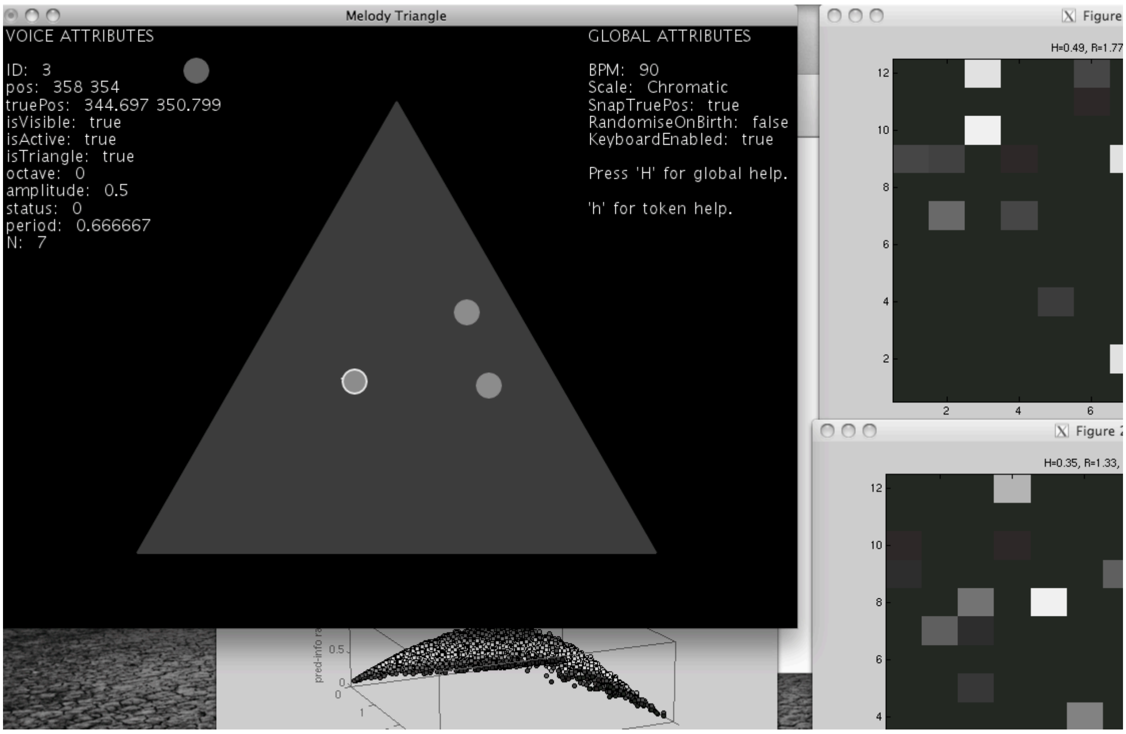


Figure 3.7: Screenshot of the Melody Triangle UI. On the right current transition matrices being played are displayed.

3.5 The Desktop Application

In the screen based interface, a number of tokens, each representing a sonification stream or 'voice', can be dragged in and around the triangle with the mouse. As in the installation, a sequence of symbols is sampled using the corresponding transition matrix, which are then mapped to notes of a scale or percussive sounds. Keyboard commands give control over other musical parameters such as the pitch register, volume, scale, inter-onset interval and instrument for each voice. The system is capable of generating quite intricate musical textures when multiple tokens are in the triangle. The overlapping and interweaving of melodies of varying periodicites and predictability is well suited for making content that could stylistically be characterised as 'minimalism', not so unlike the work of Steve Reich or Philip Glass.

This interface is quite unlike other computer aided composition tools or programming environments, as here the composer exercises control at the abstract level of information-dynamic properties.

[10]This installation was presented at Digital Shoreditch, the Bradford Science Festival, as well as the Brighton Science Festival

3.5.1 User trials with the Melody Triangle

A pilot study with six participants was carried. The participants were asked to use a simplified form of the user interface (a single controllable token, and no rhythmic, registral or timbral controls) were carried out under two conditions: one where a single sequence was sonified under user control, and another where an additional sequence was sonified in a different register, as if generated by a fixed invisible token in one of four regions of the triangle. In addition, subjects were asked to press a key if they 'liked' what they were hearing.

The hypothesis was that users would linger longer in areas of the triangle that would produce more aesthetically desirable sequences, and these would tend to be the in the areas of the triangle that are of high predictive information rate, that is, areas along the middle and lower edge of the triangle.

The subjects' behaviour were recorded, as well as points which they marked with a key press.

3.5.2 Results

Some results for four of the subjects are shown in fig. 3.8. It was no possible to detect any systematic across-subject preference for any particular region of the triangle.

Comments collected from the subjects suggest that characteristics of the patterns were readily apparent to most: several noticed the main organisation of the triangle, with repetitive notes at the top, cyclic patterns along one edge, and unpredictable notes towards the opposite corner. Some described their systematic exploration of the space. Two felt that the right side was 'more controllable' than the left (a consequence of their ability to return to a particular distinctive pattern and recognise it as one heard previously). Two reported that they became bored towards the end, but another felt there wasn't enough time to 'hear out' the patterns properly. One subject did not 'enjoy' the patterns in the lower region, but another said the lower central regions were more 'melodic' and 'interesting'.

3.5.3 Discussion

The initial hypothesis, that subjects would linger longer in regions of the triangle that produced aesthetically preferable sequences, and that this would tend to be towards the centre line of the triangle for all subjects, was not confirmed. However the subjects did seem to exhibit distinct kinds of exploratory behaviour. It is possible that the design of the experiment encouraged an initial exploration of the space (sometimes very systematic, as for subject (c)) aimed at *under-*

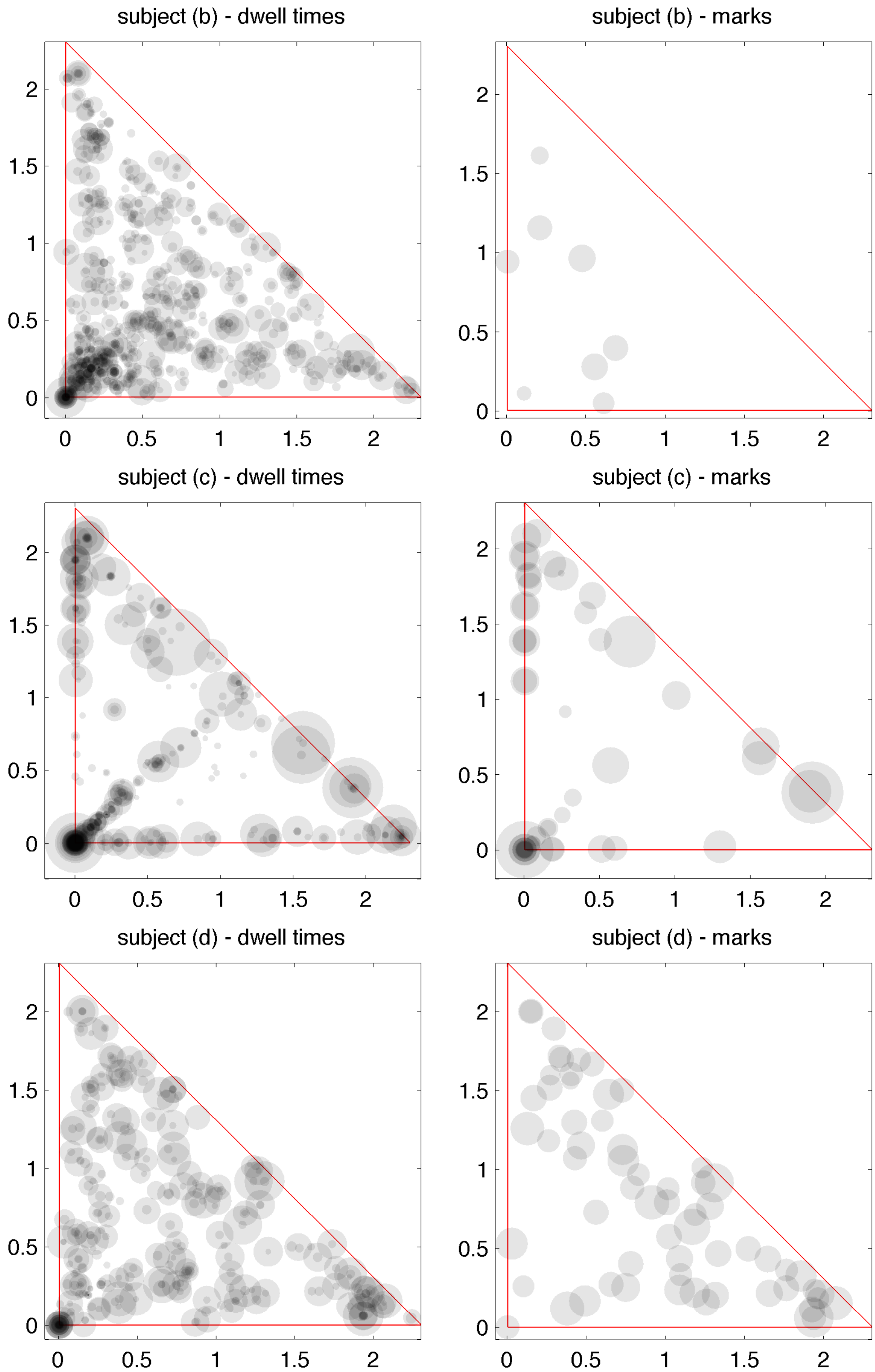


Figure 3.8: Dwell times and mark positions from user trials with the on-screen *Melody Triangle* interface, for four subjects. The left-hand column shows the positions in a 2D information space (entropy rate vs redundancy in bits) where each spent their time; the area of each circle is proportional to the time spent there. The right-hand column shows point which subjects 'liked'; the area of the circles here is proportional to the duration spent at that point before the point was marked. The plots for all users can be seen in the appendix.

standing how the system works, rather than finding musical patterns. It is also possible that the system encourages users to create musically interesting output by *moving the token*, rather than finding a particular spot in the triangle which produces a musically interesting sequence by itself.

In light of this it was clear that a different approach was required to be able to determine what kinds of melodies people prefer. One issue was that the restricted interface (one single moveable token with no additional musical controls) made for an experience of limited musical enjoyment. Additionally it was clear from the collected data that users need some time to get beyond the initial exploratory phase before it is possible to get a sense of their aesthetic preferences. This is what motivated the implementation of the mobile app version of the triangle. The app puts no restriction on time for the users to familiarise themselves, has numerous features and controls to increase the variety in musical output, and the users can submit settings when they feel like. The mobile app is discussed in section 3.6.

3.5.4 Qualitative Feedback

In parallel to the pilot study, informal qualitative feedback was elicited from users of the screen interface. Here four participants were interviewed, all practicing musicians that use computers in music production or in performance. This is with a view to establish what features would be desired for any eventual further development of the interface, for instance as a VST instrument for inclusion in a standard audio production environment.

Unlike in the pilot study where participants would not know anything about the interface before hand and were asked to 'explore' with as little instructions in possible, here the potential users are first taught how to use the system. Then were given time to play and experiment, and in informal discussion feedback and criticism of the system was sought ought. As part of a broader conversation, they were asked if they could identify the different areas of the triangle, what features of the system they liked and disliked, if they could see themselves using the system as part of their musical practice, and if so how.

Some points collected include -

- The subjects were very quick to get to grips with the properties of the different areas of the triangle, and found it quite intuitive.

- The subjects reported using the periodic/predictable half of the triangle more than the unpredictable half.

- All users desired more control over the mapping of symbols to notes, and some desired the ability to map the output of the triangle to other parameters such as to the control of filters and effect parameters.

Some comments are provided here -

> "If it was a kind of VST instrument, I would use it really a lot, definitely! Because there are not that many around that make this kind of stuff. I always love if something is generative or stochastic to generate things I would not come up with, but to generate a lot of them in a short amount of time and I'm the creative catalyst that just picks them.. and then have this kind of choices to edit probabilities, I really like that."

> "What is cool is that I can make multiple loops and they all have different characteristics and I don't have to adjust like five numbers in different places, it's in one thing, and that's what I like most, it's kind of like a macro [interface]."

These suggest that there is potential for the *Melody Triangle* as a controller in the context of electronic music production and performance. Further development to support this more effectively, such as turning it into a VST instrument for integration with common music production packages, is the subject of further work.

3.6 The Mobile App

As mentioned in sections 3.4.2 and 3.5.3, both the interactive installation and the controlled trials with the simplified desktop version of the *Melody Triangle* were unable to provide significant insights on the relationship between information properties and aesthetic preferences. Thus another approach was taken, where through crowd-sourcing via a mobile phone version of the *Melody Triangle*, data would be collected from users world wide.

In the next section, the details of the mobile phone application are provided, including how it is able to collect data on user preferences. In section 3.6.3 the results of the analysis of the collected data are provided, and in 3.7 the findings are discussed and further work is outlined.

A demo video of the mobile app is provided as item *Ex*1 of the illustrative materials accompanying this thesis[11]. Some audio samples generated with the app are also provided as item *Ex*2

[11] The video can also be accessed online - https://youtu.be/U7w69aabkqk

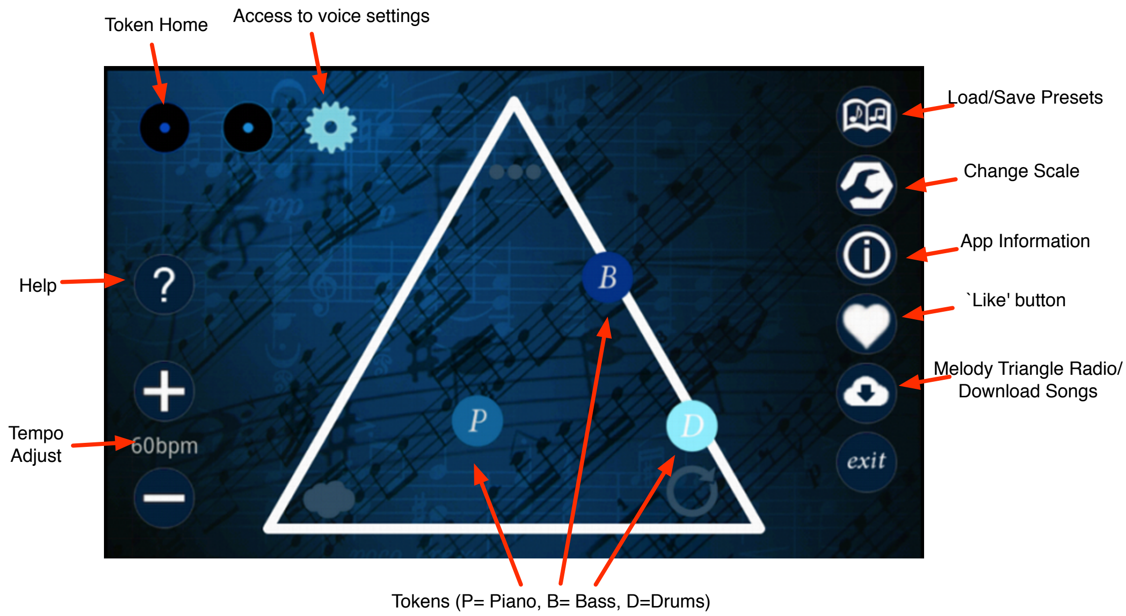


Figure 3.9: Screenshot of the Melody Triangle mobile phone application

of the accompanying illustrative materials.

3.6.1 Features

The *Melody Triangle* mobile app was developed for the Android operating system, it was launched on March 30th 2013, and is available as a free download from the Google Play app store[12]. It follows the same principles as the desktop version of the application, with the triangle the main interface element, and three tokens, that each generates a melody when placed in the triangle. An annotated screenshot of the interface is provided in fig. 3.9

When the app is first launched after install, a walk-through tutorial is provided, where each interface element in turn is highlighted and its function described with an on-screen text pop-up. At launch the three tokens, each of slightly different colours, are placed in slots on the top left of the screen. From there the user drags a token into the triangle. Once in the triangle, the melody gets generated automatically, and the vacated slot where the token was becomes a 'cog' icon that the user can press to adjust parameters specific to this token, or *voice*.

There are two global settings that the user can adjust. On the lower left, the plus and minus buttons let the user adjust the global tempo. On the right, the pressing the icon that looks like a spanner allows the user to select between three musical scales - pentatonic, diatonic and harmonic. In pentatonic mode, the transition matrices would have 6 symbols, and in diatonic or

[12]https://play.google.com/store/apps/details?id=com.qapps.melodyapp

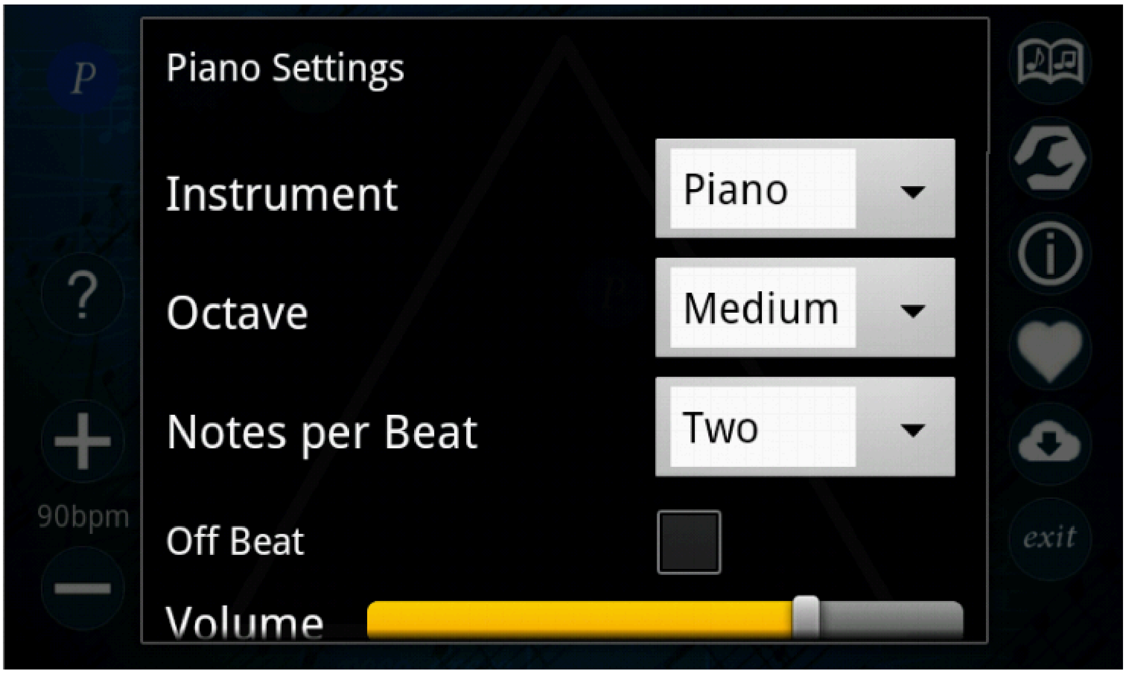


Figure 3.10: Screenshot of the menu for token/voice specific settings.

harmonic mode, 8 symbols (covering the notes of the scale).

Every time a token is moved, it is mapped to a new Markov transition matrix. Additionally a new mapping from generated symbols to sound samples is generated. The mapping was a one-to-many mapping, where different symbols could be mapped to the same note, and additionally a symbol can be assigned to a 'rest'. It is recognised that these design decisions would potentially add some noise to an analysis, however it was deemed they greatly increased the quality of the musical sequences. This would make the app be more able to sustain user engagement, a matter of critical importance in all crowdsourcing projects.

Each token can be one of three instruments, piano, bass and drums. The tokens are marked with a *P*, *B*, or *D* to correspond to the current instrument selection. Whenever a user touches the 'cog' icon that corresponds to a voice, a pop-up menu appears, allowing the user to adjust voice specific parameters. A screen shot of this menu is provided in fig. 3.10.

From there a user can change what instrument the voice is mapped to, the register of the instrument, and how many notes per-beat the voice should generate, if the token should go onto the off-beat, and a volume specific to this voice. There are three registers for the piano (high, medium, low), two for bass (high, low), and none for the drums[13]. As such a wide array of features is provided, allowing for a large gamut of possible musical outputs.

[13]Drums would always be mapped to 8 symbol transition matrices.

Table 3.2 shows the default settings of the app. These were chosen to help ensure as immediate an engagement as possible. For instance it was noted how some users of prototype versions of the app who would drag a ‘bass’ token into the triangle, but as users often would not have headphones on, the sound heard would be very quiet. As such it was decided to not have bass be one of the default instruments.

Table 3.2: Default settings of the Melody Triangle mobile app

Global BPM	90
Scale	Harmonic
Token 1	Piano, high register, two notes per beat
Token 2	Piano, middle register, one note per beat
Token 3	Drums, two notes per beat

Data Collection

The heart shaped ‘like’ button at the right hand side of the interface is the means by which users would submit their settings. Whenever it was pressed, a snapshot of the current state of the app would be taken, and uploaded to a central server. The ‘like’ button was pointed out during the initial walk-through tutorial, and users would be periodically reminded of it while playing with the app. It is these submitted settings that will be analysed in the following section.

Each submitted setting (or ‘song’) is given a unique id. This id can be entered by any user of the app to download another user’s submitted settings. This feature is available through the button with the ‘cloud’ icon.

This is also the access point to the *Melody Triangle Radio*, which allows users to listen to other users’ submitted settings easily. Upon entering ‘radio’ mode, the app queries the server for a randomly selected submission (or ‘song’), and sets the app to these settings. Additionally, a new icon appears on screen, a ‘next’ button. Whenever the ‘next’ button is pressed, another randomly selected song is downloaded and played back on the phone.

In this mode the user can quickly sample many submissions from other users. The app notifies the user that if they enjoy what they hear, they can press the ‘like’ button. This is then registered on the server, and the settings accumulate a ranking, forming a music chart, tracking the most popular songs. The chart was displayed on the web-site associated with the project[14]. Recordings of the five most popular songs are provided in the supplementary materials accompanying this thesis.

[14] http://melodytriangle.eecs.qmul.ac.uk/

Making this app was very challenging, in particular with regards to getting reasonable performance out of Android's audio libraries. The next section details how this was achieved.

3.6.2 Implementation Details

Playing a Sound on Android: Not So Easy

There are a number of different options for playing sound files on Android. Broadly speaking there are the solutions based on the native Java Android platform, options involving the underlying C platform, and third party applications and ports (libpd - *PureData* for Android, *OpenFrameworks*).

Nearly all options were explored in the development of this app. The requirements were quite modest; the app needed to be able to play sequences of short (1 second) samples to make melodies, and to have a decent level of musicality, there needed to be good timing accuracy. Additionally the samples needed to be small to limit the download size of the app while still having acceptable sound quality.

On paper this should have been straight forward; the Android platform boasts a sound file playing Java class, *Soundpool*, that is suppose to provide 'low-latency playback' of both OGG and MP3 files. However, it was quickly discovered that the Soundpool class has excessive latency, and worst of all, the amount of latency is unpredictable; sometimes the sound would be played right away, sometimes the sounds would be played up to a third of a second later. If the latency were constant, this would not have been an issue.

The other top level Java class, *MediaPlayer*, has large overheads and was not suited to playing many small samples.

The next option was then to implement a solution using a lower-level class - *AudioTrack* (SoundPool and MediaPlayer are wrappers around AudioTrack). An implementation at this level would have allowed for low-latency audio. AudioTrack is what is most often used in serious games for Android. However this turned out not to be practical because the Android API does not allow access to an OGG or MP3 decoder. Thus to make the download size of the app acceptable would have necessitated developing a custom OGG or MP3 decoder, not an option due to time restrictions.

The next option was to use *libpd*, a port of the open source Pure Data (PD) platform for Android. This is a node based programming language that is designed for audio. However the timing inaccuracies of the Java layer were inherited by PD. It was then attempted to develop a sample-accurate sequencer in PD using digital signal processing, with phasors controling the playback of samples. This was very challenging, and it took weeks to achieve the desired logic. However when the solution was finally put on the phone it was discovered that it ended up using too much CPU, causing the user interface to become sluggish and unusable.

The final solution involved implementing a ‘hack’ on the SoundPool class. Even though SoundPool was unable to play sounds on time, it turns out it can ‘un-pause’ a sound on time. Thus a sound would be ‘loaded’ by having calls to play followed immediately by a call to pause it. Then when the sound is needed, a call to un-pause or ‘release’ it would be made. However as sometimes the SoundPool class would be slow to get a sound paused, a quarter of a second of silence had to be added to the beginning of each sound file. This implementation involved multiple threads (one for each melody), each would synchronise with the others by keeping track of the system clock and the duration of its operations, and check these against the time it was scheduled to play a note.

Additionally some further restrictions to the audio content had to be made; to reduce CPU and memory load, the audio files had to be mono, low but constant bit rate OGG Vorbis files. The SoundPool class would often crash if it had calls to play files come too quickly in succession (even through that is what it was designed for), thus a semaphore had to be placed around the class so that it would never be overloaded by concurrent instructions from the different threads.

This solution yielded acceptable performance on 5 out of 6 tested devices. However if the beats-per-minute was high and the app was made to play many notes, invariably the timing accuracy would disappear. The melody playback threads were able to detect if they are behind schedule, and so when this happens, the app gradually lowers the master tempo until the phone can keep accurate timing.

A key lesson learned from developing this app is that the Android platform is not well suited to audio applications, and iOS is almost certainly a better alternative for audio intensive applications[15].

Implementing the Information Dynamics Engine

The *Melody Triangle* is based on some quite complex matrix based mathematics. The original desktop based implementation was developed in a combination of OpenFrameworks, Matlab and Prolog. Initial attempts to re-implement the stack natively on Android was hampered by a lack of

[15]However there was one possible solution which had not been tried; OpenFrameworks for Android.

suitable libraries for matrix operations in Java. Additionally the desktop version takes about two minutes on a powerful laptop to do all the necessary calculations to populate the triangle with transition matrices, and in a mobile app this would have taken an unacceptable amount of time. Thus a different approach had to be found.

The solution used the desktop implementation to ‘harvest’ all possible outputs of the information dynamics engine (the outputs are in the form of matrices, representing Markov chains). Scripts were written that would run a batch process on the desktop application, finding the output matrix for every screen coordinate of the triangle. These would then be stored into two large data files (one mapping screen coordinates to a matrix id, the second mapping matrix ids to the contents of the matrix). These data files would then be uploaded onto the phone, wherefrom a lookup of the right output matrix could be found given a screen coordinate.

However memory management was an issue. Initially the lookup tables were just loaded as arrays into memory, however this ended up taking up too much of the phone’s heap memory, causing the rest of the app to become sluggish. The solution came in using *RandomAccessFiles*, whereby the data would remain on the file storage media, but the appropriate content could be retrieved from them by calculating an offset in bytes with which to do a lookup for the appropriate information. This allowed the data to be accessible quickly without taking up the phone’s precious heap memory.

3.6.3 Results

General statistics of the app as of June 17 2015 are provided in table 3.3

Table 3.3: Melody Triangle app general statistics

App installs	3945
Submitted entries	339
Radio Listens	2663
Radio Likes	217

As can be seen, the app had a reasonable number of installs, 3945. However, the conversion rate to user submissions is unfortunately much lower, with only 339 submitted entries. Nevertheless some patterns can be gleamed from the submissions.

General statistics on the submitted entries are provided in table 3.4.

Table 3.4: Melody Triangle submissions, general properties

Submission by number of voices		**Submission by scale**		**Voices by Instrument**	
One voice	28	**Pentatonic (6 symbols)**	61	**Piano**	438
Two voice	40	**Diatonic (8 symbols)**	65	**Bass**	202
Three voice	271	**Harmonic (8 symbols, default)**	213	**Drums**	281
				Total voices	921

Referring back to the default settings in table 3.2, the following observations can be made:

- *Observation 1*: Users tend to stick to default settings.
- *Observation 2*: Users prefer to use all three voices.

With regards to tempo of the submissions, the distributions by beats-per-minute is shown in fig. 3.11.

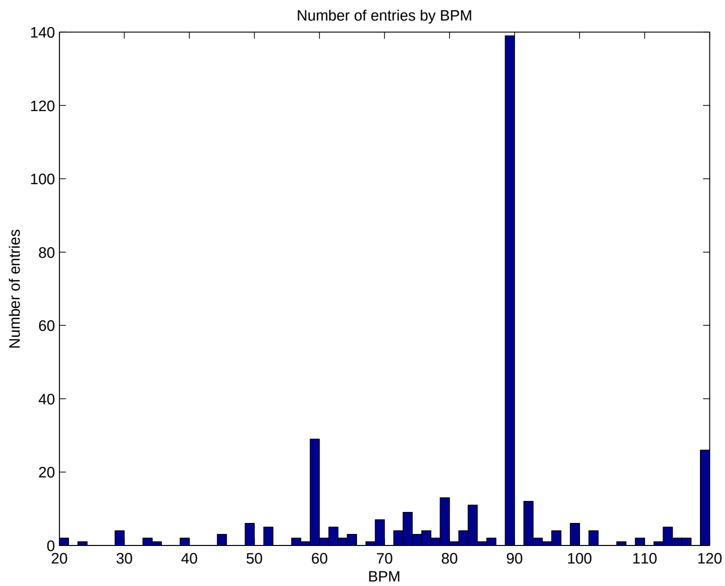


Figure 3.11: Histogram of app submissions by beats-per-minute

As can clearly be seen, by far the most submissions were at 90 BPM, which was the default tempo. This was already noted in *observation 1*. The peak at 60 BPM will be because the initial launch of the app had this as its default, this was later increased in a subsequent update to help immediate user engagment. The peak at 120 is the maximum speed of the app, and has had significantly greater number of submissions than the rest. As such the following observation can be added:

- *Observation 3*: Users like to make the app go as fast as possible.

The distribution of tokens for submissions is presented in fig. 3.12. The entropy rates and redundancy values are here normalised so that the entries that are pentatonic (with transition matrices of 6 symbols) could be placed in the same plots as those with 8 symbols. As such they represent the positions of the tokens on the interface. (Separate plots were made for the matrices of 8 and 6 symbols, but they showed effectively the same patterns).

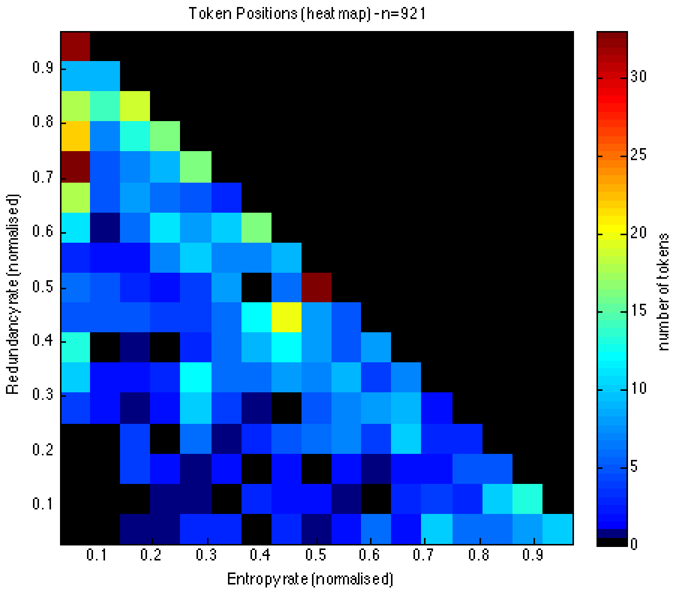


Figure 3.12: Heat-map illustrating the positions of tokens for all app submissions.

It is evident that there were many more tokens placed on the right hand side of the triangle than on the left. This would indicate an overall preference for greater predictability. It is also evident the most favoured areas are of minimum entropy and of high or maximum redundancy, that is loops of long or maximum duration. However areas 'just off' perfect long loops, just slightly off the edge, where also highly favoured.

One also sees a large peak at the middle bottom of the triangle, the position with the highest predictive information rate. There is also quite high distributions of tokens as a line down the middle of the triangle.

Additionally one notes a small clustering of tokens at the areas of maximum entropy rate, the bottom left of the triangle. Finally the last evident pattern is that areas both low in entropy rate and redundancy, areas near the top of the triangle, are clearly the least popular areas of the triangle. These are patterns that consist of very few symbols, and would approach being just one or two repeated notes.

The following observations were thus made:

- *Observation 4*: Users prefer the predictable side of the triangle significantly more than the unpredictable side.
- *Observation 5*: Users like to place token on the right hand edge, yielding perfect loops.
- *Observation 6*: Users like to place tokens in the middle of the triangle, particularly the middle bottom, the area with the highest predictive information rate.
- *Observation 7*: There is slight preference for tokens in positions of maximum entropy rate, bottom left corner of the triangle.
- *Observation 8*: Users dislike extreme repetition, the top of the triangle.

To identify if the tempo of sequences of notes had an impact on the user preference, plots of the distributions of the tokens were made with each token was waiting by how many symbols per minute it was outputting (by multiplying the global BPM by the number of tokens per beat it was generating). The heat map for this plot is shown in fig. 3.13.

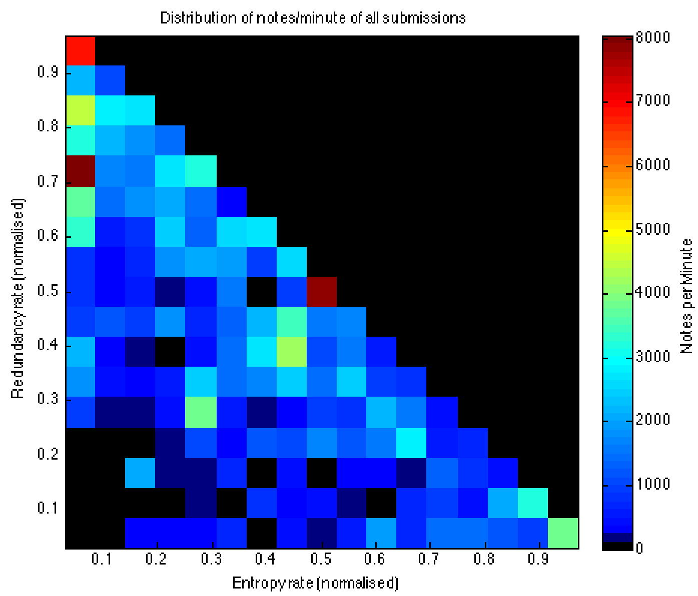


Figure 3.13: Distribution of notes per minute for all submissions

As can be seen, there is no significant difference in the overall distributions. The following observation could thus be made:

- *Observation 9*: Note rate had no discernible impact on token placement preference.

To identify if the instrument selections had an impact on placements of tokens, the positions were plotted and the average of the positions for each instrument calculated. This is shown in fig. 3.14.

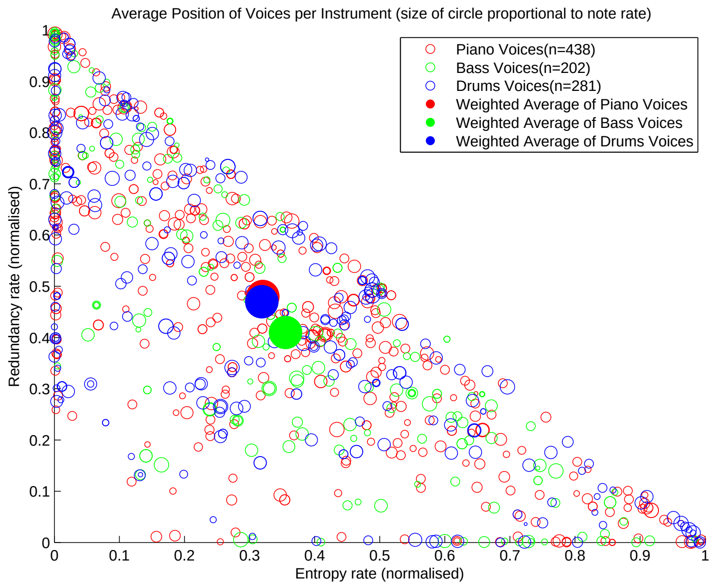


Figure 3.14: Token placements by instrument.

As can be seen, the means are very close to each other, with the bass perhaps a little more towards the unpredictability side, but not significantly. The following observation can thus be made:

- *Observation 10*: Instrument selection had no discernible impact on token placement preferences.

Melody Triangle Radio

A screenshot of the 'Melody Triangle Hall of Fame', that displays the top unloaders and the songs with the most likes in radio mode, is shown in fig. 3.15.

TOP UPLOADERS

Username	Number of Uploads
Anonymous users	196
BLUE	15
KARLYNN	14
MRSPACEMAN	12
VOUDOU	10
FFFFRED	9
KORDIKK	8

TOP SONGS

Chart Position	SongID	Composer	Recent Likes
1	TTJQXI2	Z	8
1	TAKPDIB	KARLYNN	8
2	TOEM3G8	*Anonymous user*	7
2	MW9PCSG	EX	7
2	T31FFJV	EX	7
3	M6BB97N	*Anonymous user*	6
3	TWQNOXT	*Anonymous user*	6

Figure 3.15: The Melody Triangle 'Hall of Fame' as June 17 2015.

As can be seen, the number of 'likes' is very small, with the most a song has received being 8, out of a total of 217 likes and 2663 listens. The older a song is, the longer it will have been available for listening in radio mode. As such any analysis that took into consideration the position in the charts would be skewed towards older songs (it was observed that there was little change in the positions of songs over the period of the apps lifetime).

It was thus not deemed insightful to analyse the songs based on chart position, as they would almost certainly not yield further insights. However the top 5 songs are provided for reference, as item *Ex*2 of the supplementary materials .

3.7 Discussion and Further Work

3.7.1 Success

Could the *Melody Triangle* be considered a 'success'? A metric for success was defined in section 1.2.1. It consisted of two criteria: $C1$ - *correlation between parameter choices and aesthetic value of the output*, and $C2$ - *sufficient volume of data points and significant trends*.

C1

Considering first $C1$; was the collected data really aligned to aesthetic preference? Many of the observations in this analysis perhaps do not come as much of a surprise. It would be expected that users stick to defaults, and that they prefer to use extreme settings (maximum BPM, corners of the triangle). One thing that was clear, is that there is a definite preference for the right hand side of the triangle.

As was indicated, the majority of submissions used all three tokens. As multiple voices play simultaneously their outputs will interact with each other yielding higher level forms and patterns. Even three tokens all placed along the predictable edge, as they phase in and out of each other, can yield phenomenally complex results. This increased complexity will certainly be a strong contributor to the preference for putting tokens on the predictable half of the triangle.

It is difficult to capture in the analysis the emergent phenomena of multiple tokens interacting with each other. If the app was simpler, and hence the data was more clean, then perhaps clearer patterns could have been found in the analysis. However that would have necessitated limiting the features available to the users, and in all likelihood would have decreased user engagement. It is difficult to get this balance right in such a project; on the one hand user engagement and fun are of primary concern, as without these there would be no data. On the other hand, too many features can make eliciting patterns more difficult.

The other interesting trend was *observation 6*, the tendency for users to place tokens along the middle near the bottom of the triangle; the area with the highest *predictive information rate*. The natural question is thus: was this the result of an overall aesthetic preference for these sequences? Or was this simply an effect of interface priming, and users like to place tokens there because it 'looks nice'? To address this question, it is worth considering how likely it is that such melodies would be preferred.

The *Melody Triangle* posses a fundamental limitation: the information measures are derived directly from the Markov chains, and are as such calculations based on a listener model that has the same limitations, that is, a very short-term memory. However human listeners do not have a short-term memory, and remember what has happened beyond just 'one note ago'. A sequence of high predictive information rate, from this first-order Markov model, is in all likelihood not going to be the most interesting kind of melodic sequence from a human listener's point of view. It is perhaps the most consistently *informative* kind of melody a first-order Markov chain can make, in that it constantly generates information about future likely outcomes. But as it lacks long term coherence, it could actually be perceived as rather disordered from a listener model with a greater memory.

It is thus the present author's opinion that the preference for high predictive information rate melodies is more likely due to interface priming. However to determine this for sure is the subject of further work. It is unclear how to setup such an experimental situation. As the pilot study with the single token on the desktop version of the app showed, it is not straight forward to elicit musical preferences. One could imagine designing an interface that does not have those interface priming issues; perhaps a circular interface. However it is unclear how the relationships to the quality dimensions could be maintained, and the usability of the interface (and probably user engagement) would be likely to suffer as a consequence. It would also be a clear improvement to apply a more sophisticated listener model that could take into consideration more past events. To create a *Melody Triangle* that operates on more long term models, such as a higher-order Markov processes is the subject of further work.

Attribution

To what extent can it be said that the an uploaded set of settings are really the creative work of users of the app? Although they had numerous parameters to pick from with easily understood consequences (e.g. instruments, tempo, volume), the primary controls – the position within the triangle – is an abstract notion of predictability; the user's actions are quite removed from the final audible content (see 3.4). The *Melody Triangle* does not even allow users to make a 'composition', with a start, middle and and end, instead it just lets users specify settings for never-ending stochastic processes.

There is a tension between user control and engagement. If the users could have had more control, for instance if they could specify the mapping from symbols to notes manually (rather than have them be selected by the app automatically), the parameter-space would be much larger.

This would have knock-on effects for usability and complexity, and would result in a steeper learning curve; a possible barrier to user engagement. However, giving the users greater control would give them more say over the sonic properties of the generated output. This could in turn give the users a sense of ownership over the output, which could in turn could translate to a higher level of engagement, as well as aligning their outputs more closely to their aesthetic preferences. This could make it more likely that they would share their generated works through the app, helping along the crowdsourcing of the work.

Overall the evidence suggests that criteria $C1$ was not met. There was clearly a tendency to stick to defaults or take settings to extremes, as well as there being interface priming issues with regards to the triangular interface. Additionally it is possible that there was not enough leeway for the participants to make works that really are 'their own', and hence reflect their aesthetic preferences.

C2

With regards to success criteria $C2$ – wether there were enough data points to relative to the size of the parameter space – this criteria was not met either. The number of submissions was lower than expected (only 339). This indicates that the app had some barriers to user engagement. For a user to make a contribution a significant number of hurdles need to be overcome. First, they need to come across the app, which is no mean feat in the crowded ecosystem of mobile phone apps all competing for attention. They then need to have a compatible phone, download it, and then be able and willing to engage with it and spend enough time with it to make sense of it all. Finally, they must feel that they have something that they would like to contribute. Each of these hurdles will have filtered out a large number of participants.

Medium specificity

There are some issues with medium-specificity in this experiment; there is no guarantee that the users of the app are listening to the outputs at the same volume for instance. It is well known that the perceived frequency response changes with volume(Suzuki & Takeshima, 2004); low frequencies seem quieter at low playback volumes, and appear more prominent at louder volumes. This could partially be addressed by compensating for this difference; i.e. increasing low-frequencies when playback is low volume. However before this solution could be applied, there are more difficulties at hand. For a start, the frequency response is determined by the make and model of phone (of which there are thousands for Android), as well as the make and model of headphones. Further some participants could be using the app without headphones and only using the speakers from the phone. Even if the frequency response could somehow be evened-out across participants, it is expected that the overall appreciation of a piece of music would be in some form affected by the loudness of playback.

These are effects that could skew any results gathered. In a future study, these issues of medium-specificity could be addressed by seeking to normalise as many of the above issues as possible. For instance, the app could be deployed only for one kind of smartphone, such as Apple's iPhone. This would reduce differences in frequency responses in the playback. It would also increase the likelihood of users having the same headphones (as many iPhones ship with headphones). Additionally, the app could instruct the users to use their headphones before beginning, to prevent users from working with the app using the phone's speakers, and further could instruct the users to set a specific volume. However even if these variables were controlled in this manner, there are still countless external variables that come into play that would affect the decisions made. A participant playing with the app on a loud and crowded train will interact differently than if they were in a quiet comfortable environment. However, given enough participants and submissions these difference could average out. If there were hundreds of thousands of submissions, preferences and clustering of data would occur with respect to the 'average loudness' and the 'average frequency response' of playback. However, in this case there was nowhere near enough submissions for such an averaging to be possible, and the variability of the presentation conditions will have introduced significant noise to the gathered data.

Having failed to meet $C1$ and $C2$, the *Melody Triangle* as an experiment in empirical aesthetics did not succeed. In other ways however the *Melody Triangle* shows promise. The desktop application shows good potential as a performance tool for electronic music or as an idea generator in composition; further developing it the subject of further work. Additionally the mobile app has demonstrably yielded a significant number of pleasant and interesting pieces of music[16], as can be heard in the supplementary materials provided.

The *Melody Triangle* provides a new representation; a particular and unique way of conceiving and understanding sequences of musical events. The conceptual space that the *Melody*

[16]As well as receiving a good overall ratings on the Google Play App store; 4/5 stars

Triangle leverages is in some sense a (modest) ‘discovery’. Just as the discovery of the colour circle allowed for reasoning about colour in terms of cyclical space, the triangle affords reasoning about musical predictability in terms of its geometry.

3.7.2 Future Work and Improvements

Improving Engagement, Usability and Usefulness

What changes could have been made it to make this experiment a success? The primary issue has been one of user engagement; there were simply not enough submissions to be able to begin to uncover the underlying geography of the fitness landscape. Currently the app is in a middle ground, it is does not offer fine-grained enough control to be a genuinely useful musical interface for a hobbyist or professional musicians. Conversely it is also perhaps a bit too complex and esoteric for the average non-musician to fully understand and make use of. A future implementation should target either explicitly general crowds, and make the app more game-like, fun and easier to use, or conversely it should target musicians and music makers, offering fine-grains controlled over symbol to note mappings, and provide features that would allow it be used in a performative or recording context, such as being able to output MIDI or OSC to integrate the app with other studio tools. If the app became a genuinely useful musical tool, then it would be more likely to be able to reach a critical mass of use, and only then could enough data be gathered. It could also take the form of a VST/AU instrument for incorporation into common DAWs. Greater, more explicit control over the musical outcome could make the users feel that they had greater ownership over their work, and be thus more willing to share it.

More Sophisticated Listener Model

An area of further work is to apply a more sophisticated listener model to the triangle. First-order Markov processes, with their very short term ‘memory’ are convenient because it is possible to extract exact information measures of redundancy and entropy rate. However humans have a more sophisticated long term perception of musical events, and as such a more sophisticated model could more accurately represent human perception of predictability and surprise. Further research is needed to determine what kind of information measures are needed for a more sophisticated listener model, such as one based on higher-order Markov processes. Wether it is possible to flatten these out into quality dimensions for the creation of a musical interface remains to be seen.

Alternative Interfaces, Data Gathering Methods

The triangular interface appeared to have issues of interface priming, and as such was likely polluting the gathered data. Alternative methods for gathering preferences with respect to information measures are needed if a map between these measures and aesthetic preferences is to be built. This could take the form of alternative interfaces free of priming (perhaps some sort of circular interface), or even A/B preference comparisons between pre-generated sequences might be more effective in gathering this data.

Application to Other Domains

An intriguing area of possible further work is to apply the *Melody Triangle* to other modalities. The triangle itself operates in the abstract, and as such could be applied to any domain, such as sequences of colours in video, or flashing lights. An exciting area of possible exploration is to apply it to the level of really short musical events, the domain of 'microsound'(Roads, 2004). It could be used to control a granular synth for instance, rapidly triggering tiny snippets of audio to build sonic textures. Markov chain based granular synths have been explored previously(Miranda & Junior, 2005). It would be intriguing to hear what the phenomenal properties of the different areas of the predictability conceptual space correspond to in the timbral domain.

In this chapter the *Melody Triangle*, a musical interface derived from information theoretic properties of Markov processes, was presented. Its theoretical basis was outlined, including how it provides a mapping to a *conceptual space* of predictability. The *Melody Triangle*'s three incarnations were detailed; an interactive installation, a desktop application and a mobile phone application. The interactive installation functioned well in its role as a communicator of research concepts to the wider public, however it was not practical for eliciting aesthetic preferences. The desktop version of the app, although demonstrating potential as useful performance tool and composition aid, studies carried out with it in controlled contexts were also unable to identify patterns in aesthetic preferences. The mobile phone application allows users world-wide to submit their favoured settings, with compositions of notable worth uploaded by users of the app. Some overall trends were identified, in particular a clear preferences for patterns of increased predictability. Additional identified trends however seem to be subject to issues of interface priming, and limits to user engagement were identified. This prompted the question of how these could be mitigated, and some suggestions for this were presented as further work.

But could it be possible to bypass the issues of interface priming, by removing the interface from the picture entirely? Is it possible to not have computational generative process mediated by layers upon layers of mappings and devices, as the *Melody Triangle* is, but to be intimately embodied and explored in a natural way? Could the interface be made invisible, or even non-existent? These question inspired the creation of a novel system and art project, the subject of the next chapter: *Keyebernates*.

Chapter 4

Keyebernates

In chapter 2, some of the difficulties and challenges of parameter search in algorithmic composition and design were outlined. One of the issues is that parameter-spaces are often vast, making an exhaustive search of the space not practical. Further little can be predicted about the outcome of a generative process from parameter values; similar parameters can yield very different results.

The *Melody Triangle* embodied one approach to facilitating the search, by adding an intermediate interface layer mapping phenomenal similarity to parametric similarity. However such an approach cannot be applied to any generative processes, as it requires extensive knowledge and theory of the generative processes *a-priori*, and more often than not, this knowledge is not at hand.

However the difficulties do not end there. In traditional modes of creation, such as using a paint brush, there is an immediate and clear causality between action and output. As the artist applies paint to a canvas, instant sensory feedback enables her to continuously make evaluations of the work-in-progress. There is an embodied *fluidity* to this the process. The final result is the emergent consequence of the *action-perception cycle* playing itself out as it loops though the designer, her actions on the canvas, and her perception of the canvas. The designers' sketches are the output of a "graphical conversation with the materials of design"(Schon & Wiggins, 1992), however in algorithmic and generative design, the free-flowing conversational nature of this interaction is difficult to maintain.

When composers and designers work with generative processes, the search through parameter-space commonly consists of discrete, discontinuous steps of parameter selection and output evaluations. In explicit generative creation, this discontinuity can manifest itself in the designer's interaction with the substrate of the generative system, for instance while waiting for a computer to execute some code to render some media, or in the case of manual generative design, carrying out the procedures until it is possible to see the emergent results. Only then is the artist able to pass judgment on the work, and evaluate the parameter-selections or the design of the generative process itself.

This chapter introduces *Keyebernates*, an experimental system for control and navigation of the possibilities afforded by visual generative systems. *Keyebernates* was conceived to addresses the lack of fluidity and embodied control that is feature of common generative design methods. By using eye-tracking and gaze points as the modality of input, users navigate the parameter-space of a visual generative system with their eyes, the parameter values most gazed upon being reinforced, in a continuous and reactive navigation of parameter and solution space. When a viewer observes *Keyebernates*, they do not directly see changes as they happen. Instead by using the real-time data from an eye tracker, *Keyebernates* ensures that what is attended to does not change, and rather all change happens slowly and subtly in areas of the screen not currently attended to.

Like the other practical explorations in this thesis, a key aim of this study is to see if it is possible to uncover the objective, inter-personal dimension of aesthetic judgments on a generative artefact; to find the elements that are common across individuals when value judgments are made on a parametric design. If it is possible to map out the parameter-space of a visual generative process with respect to some form of objective aesthetic desirability using gaze – and thus meet the criteria for 'success' defined in 1.2.1 – it could shed light on the cognitive mechanisms involved in making aesthetic judgments that are common to all individuals. Further, such a map could inform the construction of creative systems to assist in generating visual patterns that would be judged to be of aesthetic value by most people.

Keyebernates is both an experiment in HCI and experimental aesthetics, but also is an art work in itself. Drawing explicitly on the metaphor of *kybernetes* - the steersman steering a ship across the chaotic forces of currents, from which the field of cybernetics draws its name - this experimental system interprets the perceiver's gaze as a steering force against randomising chaotic forces.

A brief survey of the literature on the relationship between gaze and interest is provided in

4.1.1. This is followed in 4.1.2 by an overview of previous work where eye tracking is used as control modality.

In section 4.2 the architecture of *Keyebernates* is outlined, describing in detail how the viewer's gaze is used to navigate parameter and solution space. Section 4.3 then describes a user study with *Keyebernates*, where it was used with a very simple visual generative process consisting of overlapping circles.

The study was divided into two parts. In the first part, participants passively observed the screen, implicitly guiding the search through parameter space. As will be shown the users tended to pull the system towards similar areas of parameter space, suggesting that even within this very simple generative process, certain areas of the space yielded more visually compelling outputs then others.

In the second part of the study, participants were asked to actively search for a particular pattern. This was to test if *volitional* directed gaze, could be used to navigate towards appropriate areas of solution space. The results indicate that users are able to 'will' *Keyebernates* towards particular parameters with their gaze, however only across small distances in parameter space. As will be discussed, this is because the design of *Keyebernates* favours a slow detailed exploration of parameter space, and does not easily afford traversing large parametric distances.

In section 4.4 the results of the study are discussed, directions for further work are presented, and *Keyebernates*'s place and significance in a broader theoretical context is outlined.

4.1 Context

4.1.1 Gaze and Interest

Human eye movements have been the subject of research since the late 1800s(Rayner, 1998). This early research identified many of the fundamental eye movement facts, such as the size of effective vision (about 1°to 3°), and characteristics of saccadic eye movements (the quick movements of the eyes) such as their latency and angular distances. Humans eyes, unlike a camera, do not gaze upon a scene in a fixed steadiness, rather the eyes move and jump around to dynamically construct a mental representation of the scene.

Gaze is understood to be "both simultaneously bottom-up, stimulus-driven as well as top-down, goal-oriented"(Duchowski, 2002). This distinction demarcates what is called *non-volitional* from *volitional* gaze. Top-down goal orientated aspect of visions was already extensively studied by Yarbus in 1967. He identified that when examining complex objects, such as a painting, gaze is fixated more on certain areas of the object over others and suggests that gaze can be read to elucidate internal cognitive mechanisms ("Eye movements reflect the human thought process"(Yarbus, 1967, p. 190)). In particular, he identified that the task at hand plays a fundamental role: "depending on the character of the information he must obtain, the distribution of the points of fixation on an object will vary correspondingly, because different items of information are usually localized in different parts of an object"(Yarbus, 1967, p. 190).

However this is perhaps not so straight forward, as Duchowski explicates

> "When examining a scanpath over a visual stimulus, we can often say that specific regions were looked at, perhaps even fixated (following analysis of eye movements), however, we can not be fully confident that these specific regions were fully perceived. There is (currently) no simple way of telling what the brain is doing during a particular visual scan of the scene"(Duchowski, 2002).

Although we are a long way from having a complete understanding of the biological mechanisms of the human vision system (see(Cavanagh, 2011) for a good overview of the field), there has been significant progress in modelling certain aspects of visual scene attention.

Itty and Coch outline a "biologically plausible computational modelling of a saliency-based form of focal bottom-up attention"(Itti, 2001). They suggest that bottom-up visual attention to a stimulus is highly contingent on surrounding context and propose that for a visual scene, their 'saliency map'(Koch & Ullman, 1985), that encodes the visual 'conspicuity' of a stimulus in a scene, is a plausible model for an aspect focal visual attention control strategy (for instance a red ball in a green field attracts gaze). 'Inhibition of return' where a stimuli previously attended to are prevented from being attended to again is another mechanism at play, and scene understanding and object recognition also are factors in how the gaze is directed in scene attention(Itti, 2001). Gaze is not simply directed to where things are interesting or pretty.

Numerous studies of gaze patterns during the contemplation of artworks have been carried out. Rather than being able to establish a link between gaze and aesthetic quality, these studies provide an insight into perceptual mechanisms at play in object and scene contemplation (e.g. (Buswell, 1935; Yarbus, 1967; Krupinski, Locher, Nodine, & Mello-Thoms, 2007; Wooding, 2002)). As a viewer perceives a painting, after an initial short term rapid holistic impression is made, where the viewer gets a 'gist' of the picture, individual areas of the image are then attended to in greater detail(Krupinski et al., 2007). Statistical techniques can be applied to the gaze patterns to make a ‘fixation map’ demarcating the ‘regions of interest’ in the work(Wooding, 2002; Santella & DeCarlo, 2004). Where those regions are depend on many factors and vary from individual to individual, however it has been noted that there is often a link between the number and duration of fixations in regions of the artwork, and how detailed the those regions are(Wooding, 2002; Buswell, 1935).

A significant body of research suggests that there is a correlation between interest and gaze fixations, both in task orientated studies, and in passive observation. Pieters and Warlop(Pieters & Warlop, 1999) have found strong correlation between gaze fixation and saccadic movements with preference selections of branded items. Plumhoff and Schirillo(Plumhoff & Schirillo, 2009) similarly found that gaze fixation durations and saccade-distances correlated with aesthetic preference of paintings. Holmes and Zanker(Holmes & Zanker, 2012) observed the gaze patterns when viewers are presented with 2 to 8 images in a circular arrangement, and found that the accumulated fixation duration correlated strongly with preference. Additional feature of the gaze data, such as the number of times the gaze returns to the object also correlated with aesthetic preference. They then combined these features to create an ‘oculomotor signature’ for the prediction of stimuli preference.

Research on eye tracking in image search tasks suggests that information in gaze patterns can be used to make inferences about human interests. Haji Mirza et al.(Haji Mirza, Proulx, & Izquierdo, 2012) extracted a large number of gaze-data features in an image search task, and using statistical techniques extracted weightings for the features that provided the best prediction of interest. They found that the most significant features related to the durations of gaze visits of the stimulus (e.g total visit length, average duration of visits). Hardoon and Psupa(Hardoon & Pasupa, 2010)successfully applied similar techniques in an image search system supported by implicit gaze information.

The earlier research on gaze used complex analogue mechanisms and manual techniques to collect gaze information. This was then superseded by computer based eye-tracking technology, which has became more accurate and affordable in recent decades. Today eye-trackers are no longer just used exclusively in visual perception research, but as interface mechanisms. These are briefly surveyed in the next section.

4.1.2 Eye-tracking as Control Interface

Eye tracking systems are commonly employed in a wide variety of disciplines. Their application in research range from the study of visual perception and medical research, to how gaze behaviour or interact with environments, interfaces, media, or even other humans. Eye tracking data is often used as feedback to inform the designs of user interfaces, web pages, advertisements and products on the supermarket shelf(Duchowski, 2007). However their application as an interface modality is less common, yet has gained increased interest in recent decades as researches try to leverage the real-time indicators of overt attention that the technology provides.

One of the first applications of eye-trackers as interactive modality was presented by Jacob(R. J. K. Jacob, 1990) with a gaze-based informational display where interface elements would be activate by gaze. Jacob identified the *Midas touch*[1] problem, whereby any interface element gazed upon would be activated; inhibiting an effective useful interaction.

The Midas touch is a well known issue in eye-tracking interface research, and a large amount effort is spent on mitigating it. Jacob suggests both eye blinks and dwell time as possible ways to determine if an action should be performed. Fundamentally the problem stems from the fact that the eye is a perceptual organ; to overload it with active capabilities will unsurprisingly lead to difficulties.

However eye-tracking has provided great benefits to the handicapped. Numerous systems have been developed that use gaze for eye-typing such as Dasher(Tuisku, Majaranta, Isokoski, & Räihä, 2008) (see (Majaranta & Räihä, 2002) for a survey text typing systems), as well as a drawing interface to replace the traditional mouse, such as the *EyeDraw* system developed by Hornof and Cavender(Hornof & Cavender, 2005).

Out with the context of handicap support, eye-tracking applications often use gaze not as the sole mode of interaction, but as a supplementary interface mechanism. Duchowski points out: “the eye tracker may serve better as an indirect indicator of the user’s future selective intent (i.e., serving as a mouse pointing accelerator rather than the mouse pointer itself)”(Duchowski, 2002). This approach is used by a number of systems (e.g. (Stellmach & Dachselt, 2012; Zhai, Morimoto, & Ihde, 1999)) that combine gaze with various other input modalities. Additionally gaze-tracking having been explored virtual reality contexts (e.g (Starker & Bolt, 1990; Tanriverdi & Jacob, 2000)), as well as applied to pervasive computing, where attempts to analyse eye move-

[1] After the mythological King Midas, who’s touch would turn objects into gold.

ments obtained through head mounted displays are used to provide contextual information about user's activities(Bulling, Roggen, & Troster, 2011).

A subclass of gaze research make use of what is generally termed 'gaze contingent displays', displays that react in real time and modify their contents to where users look. In such systems (e.g. (Reingold, Loschky, McConkie, & Stampe, 2003; Loschky & McConkie, 2000)), gaze alter the scene to either manage computational resources, or as tests human visual perception. *Keyebernates* could be classified as a type of gaze contingent display.

Recently the use of eye-tracking systems as creative modality have been explored. DeCarlo and Santella(DeCarlo & Santella, 2002) have used gaze information to define parameters for artistic re-renderings of photographs. Jowers et al(Jowers, Prats, McKay, & Garner, 2013), in their 'Designing with Vision'[2] research project sought to use eye tracking within the context of shape-grammar based computer aided design (CAD) system, demonstrating a system that "can select shapes and subshapes without the cognitive overload that can so easily interrupt creative flow and so stifle creative thinking"(Jowers, Prats, McKay, & Garner, 2011).

A more unusual case of creative use of eye tracking use Lasse Scherffig's *Eye-Vision-Bot*(Scherffig, 2005). An eye tracker and projector are place in an art gallery, users would simply be presented with an array of images, the ones focused on growing in size, and new images are retrieved based on the ones selected. Essentially a gaze controlled image retrieval system re-cast as a work of art.

Some researchers have also explored the use of gaze tracking as a way to define the fitness of artefacts in evolutionary algorithm contexts(Holmes & Zanker, 2012; Pallez, Collard, Baccino, & Dumercy, 2007), suggesting it as a means to overcome fatigue inherent in doing selections for interactive evolutionary computation. In such work, one of a number of discrete stimuli is selected and these selections determine their evolution. This occurs in temporally discrete steps of selection and generation.

4.2 Approach

Keyebernates seeks to enable a *continuous* navigation of a generative system's possibilities with real-time eye tracking as the feedback mechanism to 'steer' this navigation.

Instances of a class of a generative visual artefact populate the screen. The output can either

[2]http://design.open.ac.uk/DV/

consist of individual distinct generative artefacts, or can be used to evolve textures that are the result of spatially overlapping entities (see 4.2.1). *Keyebernates* counterbalances two forces; a noise source that increases variety in the parameter values, which can be understood as a process of *differentiation*[3], discussed in greater detail in section 4.2.3. Conversely the feedback from the eye tracker drives a process of *integration*[4], where the parametric variety of the output decreases, this is described in detail in section 4.2.4. This interplay between the differentiating and integrating forces determines the path traveled through parameter space.

In *Keyebernates*, the user/perceiver attends to the solution space directly, rather than selecting parameters in parameter space. There is an assumption in the design of this system that gaze fixation is related to aesthetic preference. It was noted previously that gaze patterns are complex and contingent on many factors, however as studies outlined in the section 4.1.1 show, there is strong evidence to support that assumption.

One key feature of *Keyebernates* is that the perceiver does not directly see changes as they happen. Changes in the navigation of parameter and solution space occurs only where the gaze is *not* fixated, this is made possible by the real-time nature of the gaze-tracking system. As such, change only occurs in the peripheral vision[5]. Further, all change happens in a slow continuous fashion, and is smoothly interpolated. In observing *Keyebernates* the viewer becomes aware that it changes, as when they re-visit an area of the screen that they previously attended to, it may have a different content. Ensuring that this trick works and that the changes in peripheral vision do not ‘pop out’ places a strict restriction on the speed at which the space can be explored. As will be shown, this speed restriction is an impediment to this interface being practical in a design context, and suggestions for dealing with this are discussed in section 4.4.

Keyebernates attempts to remove all the intermediate layers between the consumption of the output, and the parameter selection process. This is in contrast to the *Melody Triangle* which provides multiple layers of separation, of mapping between user control and the output of the generative system. There is the triangle interface itself, which provides a mapping to subjective predictability of the output, which in turn maps to the parameters input into the Markov process, which in turn generates the output symbols, which are mapped to sound generating processes. *Keyebernates* on the other hand collapse the distinction between consumption of the rendered

[3]In the non-mathematical meaning of the word; to increase differences

[4]Again in the non-mathematical meaning of the word; to make things more similar.

[5]This is similar to childhood game of ‘statues’, where the output ‘freezes’ upon being attended to.

output and parameter selection as much as is possible; the consumption of the output *is* the input modality.

Another fundamental distinction, is that the *Melody Triangle* is time invariant. It has no memory and does not build upon what has happened before[6]. As users pressed the 'like' button on the mobile app, the current settings were stored, however there was no sense in which the collected feedback of the users was accumulated to improve upon the system[7]. In *Keyebernates* however the perceivers continually provide implicit feedback on the value of the parameters just by consuming the output, continually building upon and steering the search for parameters.

As such, *Keyebernates* confounds the roles of the designer with that of the consumer of the work, merging them into one common role.

In the next section the kinds of the generative processes that can be used with *Keyebernates* are outlined, and details of the generative system used in the user study are provided.

4.2.1 Generative Processes

Essentially any visual generative system can be used with the system, as long as:

- The output is local to an area of screen.
- Continuous changes in parameter values correspond to continuous changes in visual output (otherwise the sudden changes will attract gaze).
- The output can be rendered in realtime.

For the studies, as simple generative system of overlapping circles was used, as matrix of 70 by 56 circles on a screen of 1280 by 1024 pixels. Constraining the experiments to a simple system allows for detailed observation and understand the dynamics of the system, however in principle any generative process that satisfies the constraints listed above can be used[8]. In fig. 4.1 are examples of the output with their associated parameter values.

Keybernates uses an internal representation of parameter values where each parameter is a floating point value ranging from 0 to 1, with the navigation beginning with all parameters set to

[6]with the exception of the one note 'state' of the Markov processes

[7]The 'Melody Triangle' radio was a first step towards this, by building crowdsourced rankings of the users feedback, and allowing users to modify the songs they liked, however as discussed, unfortunately this feature did not have very many users.

[8]Although the more parameters the generative process uses, the slower the navigation of the space.

0.5. To use a generative process with *Keybernates*, a mapping from this internal representation to the parameter space of the generative process needs to be defined.

This generative process used in the experiment takes three parameters:

- P_1 - *circle size*: the diameter of the circles. 0 maps to 3 pixels. 1 maps to 20 pixels
- P_2 - *offset in x*: horizontal displacement of the circles. 0 maps to -30 pixels. 1 maps to 30 pixels. (Multiplied by -1 for each alternating column).
- P_3 - *offset in y*: vertical displacement of the circles. 0 maps to -30 pixels. 1 maps to 30 pixels. (Multiplied by -1 for each alternating row).

The offset values, P_2 and P_3, are multiplied by -1 for each alternating row and column to cause the circles to overlap. This means that offset values equidistant from 0.5 would yield perceptually the same result. For example a P_2 value (offset in x) 0.6 will have the same visual effect as if it were set to 0.4. As can be seen, even with such a simple system there is a large gamut of visually distinct kinds of outputs. This is due to an *emergence* of higher level forms and gestalts resulting from the intersection and overlaps of the circles.

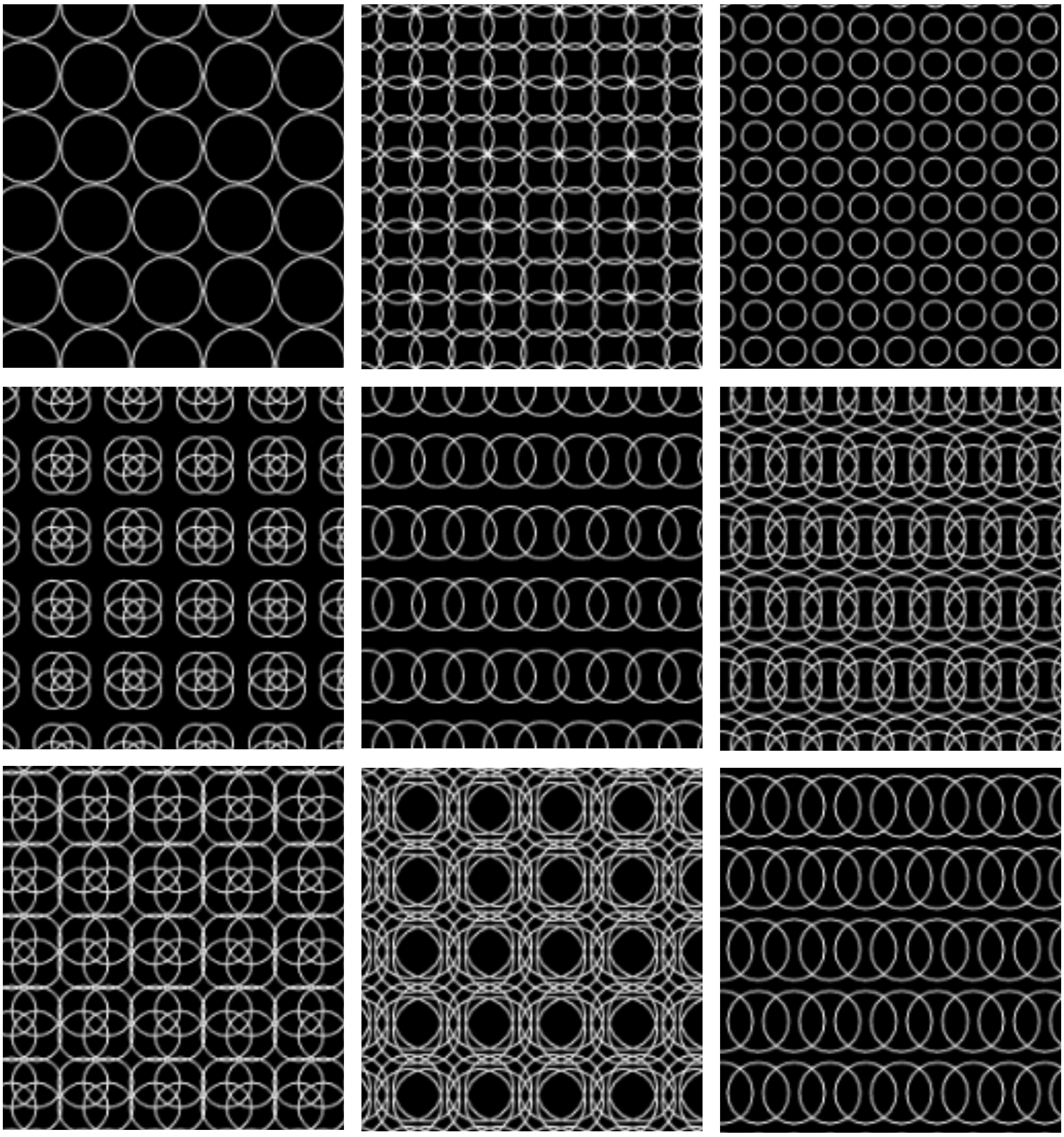


Figure 4.1: Example outputs of generative process used in the experiments. The parameter values are as follows:
Top left: circle size: 0.96, offset in x: 0.42, offset in y: 0.42
Top middle: circle size: 0.55, offset in x: 0.23, offset in y: 0.27
Top right: circle size: 0.29, offset in x: 0.24, offset in y: 0.26
Middle left: circle size: 0.44, offset in x: 0.15, offset in y: 0.33
Middle middle: circle size: 0.68, offset in x: 0.27, offset in y: 0.42
Middle right: circle size: 0.73, offset in x: 0.25, offset in y: 0.5
Bottom left: circle size: 0.64, offset in x: 0.3, offset in y: 0.19
Bottom middle: circle size: 0.82, offset in x: 0.25, offset in y: 0.14
Bottom right: circle size: 0.92, offset in x: 0.32, offset in y: 0.

4.2.2 Parameter Planes

In *Keyebernates*, each of the parameters of the generative output is represented as a two-dimensional plane, virtually existing behind the screen. These are referred to as *parameter planes*. Thus if the generative artefacts take three parameters, there are three parameter planes. For an artefact at a given screen location, its rendering parameters are drawn from the parameter planes at the same screen position (see fig. 4.2).

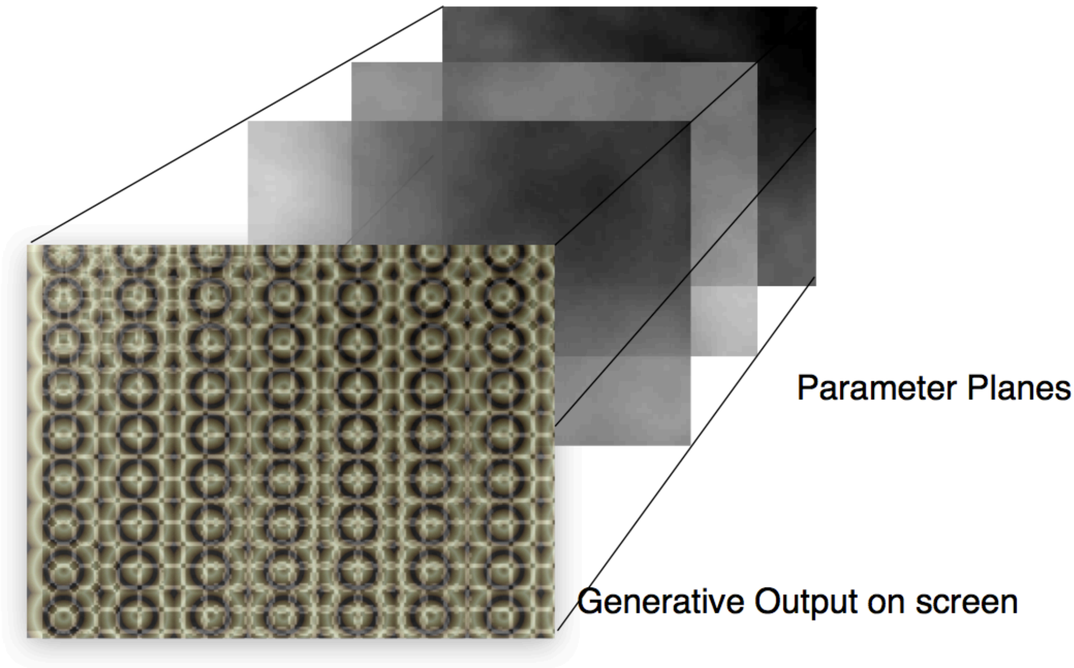


Figure 4.2: Relationship between parameter planes and output on screen. Each instance of the generative artefact on screen takes its parameter values from the parameter planes at the same coordinates.

The parameter planes contain floating point values between 0 and 1, with all values set 0.5 at the start[9]. There is then a mapping between this floating point value to the parameters used for the generative system, as described earlier. As an example fig. 4.3 shows a parameter plane, mapped to circle size and rendered as grey-scale under the corresponding circles (naturally the parameter planes are not rendered in actual use).

However, this mapping can require a scaling to discrete integers (for instance the radius of a circle is measured in pixels). In these cases discrete boundaries appear in the visual output across the population of artefacts. Since edges and sudden changes are salient and attract the gaze, a

[9]However in principle the parameters are unbounded, and it is possible to steer the system beyond this range. This never happened in the studies carried out however.

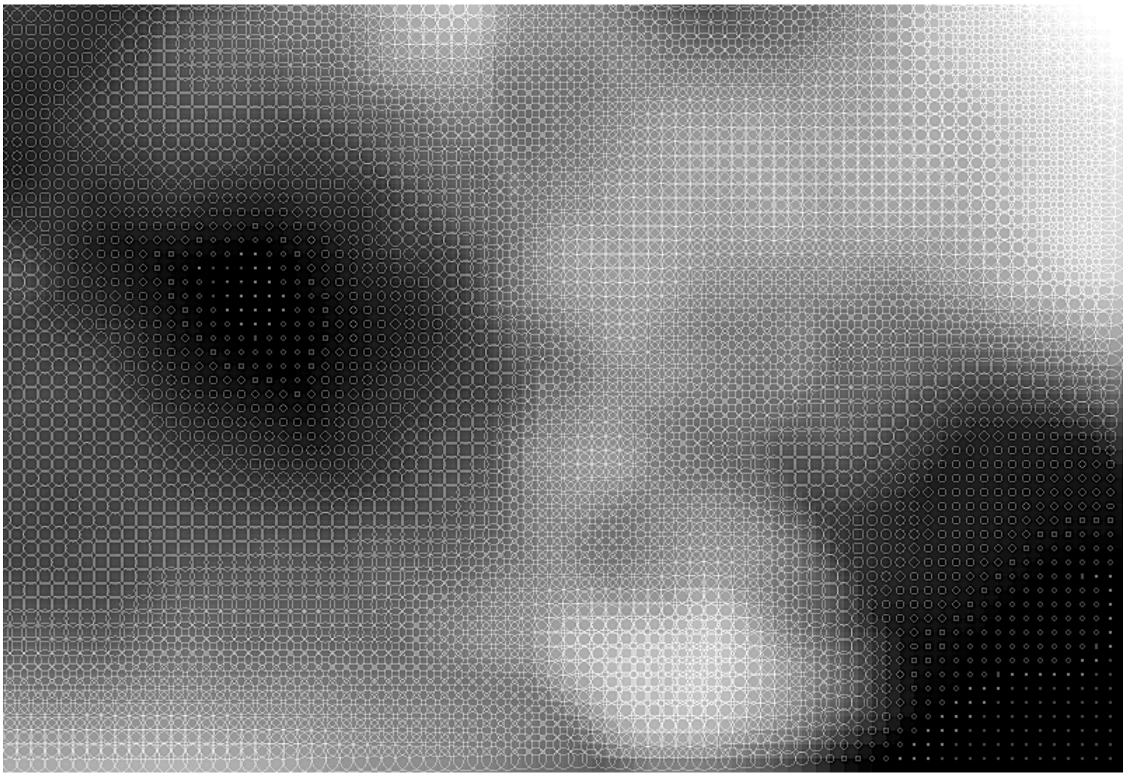


Figure 4.3: A parameter plane mapped to circle size. The parameter plane is rendered as grey-scale here for illustration, but is invisible when system is used.

temporal blur is applied to the output, turning the discrete changes into smooth transitions (see fig. 4.4).

As parameters correspond to distinct aspects of a generative output, the same amount of change in different parameters can correspond to radically different levels of change in visual similarity. To address this it is possible to apply distinct scaling functions to the mapping of the parameters. For instance one parameter plane can be made to map logarithmically to parameter values, and another linearly. When applying a generative process to *Keyebernates*, the scaling is adjusted for each parameter to ensure that the same amount of change in parameter values corresponds to roughly the same amount of perceptual change. In the simple generative process used for the experiments however, this feature was not needed.

Central to *Keyebernates* is the interaction between a randomising processes, where noise is applied to the parameter planes, and the unifying force of the gaze. The noise process increases the variety of the output on screen - a process of *differentiation* - and the gaze decreases the variety on screen, a process of *integration*, both are described below.

A video illustrating the system dynamics of *Keyebernates* is provided as item *Ex*3 of the

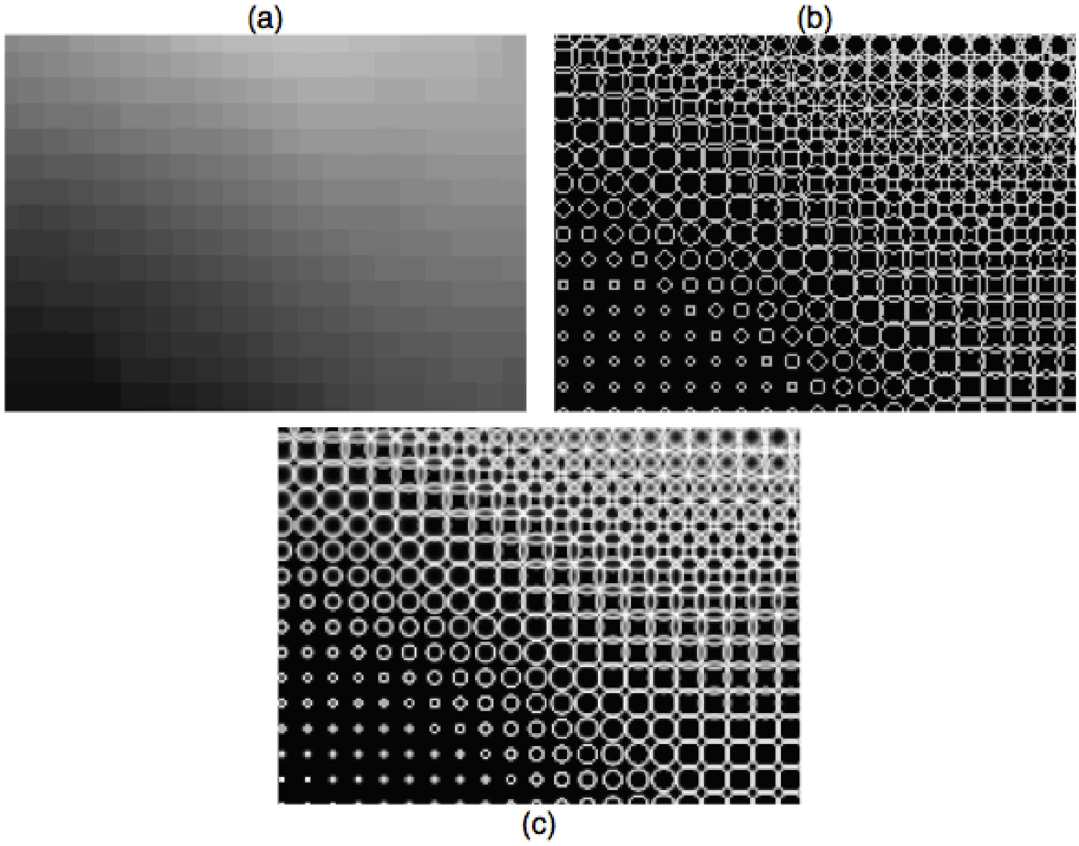


Figure 4.4: (a) Parameter values from the parameter plane, that map to circle size (b). In (c) is the result of a temporal blur (slightly exaggerated for illustrative purposes), that mask some of the effects of the quantisation of the parameter plane's values to circle size.

illustrative materials accompanying this thesis[10].

4.2.3 Differentiation

At any one time a range of parameter values is represented on screen, the values changing each frame. To ensure that that there are no sudden - and hence visually salient - changes in parameter values, three dimensional Perlin noise(Perlin & Hoffert, 1989) is used to drive the change. Perlin noise is a method of efficiently generating graphical textures. Often used to model water, smoke and clouds, it is an elegant way to provide both temporally and spatially continuous, yet seemingly random values across a plane (see fig. 4.5). Two of the dimensions are interpreted as spatial, and one as temporal, it can be understood as a video of smoothly changing values. The Perlin noise algorithm is deterministic, allowing for the possibility to replay an exact evolution and repeat experiments. However in the studies described in section 4.3, the Perlin noise

[10]The video can also be seen online https://youtu.be/1-NSAf7CGoA

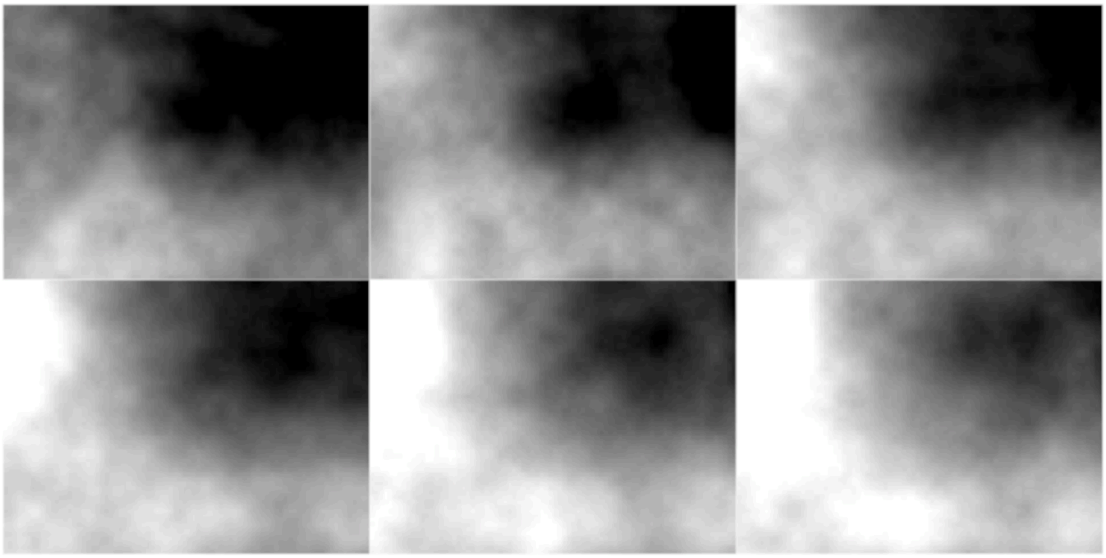


Figure 4.5: Sequence of three-dimensional Perlin noise, the third dimension is time. Note both the spatial and temporal continuity.

algorithm was seeded with random values each run.

Each frame, Perlin noise is added to the values in the parameter planes. This ensures that changes in the output are continuous and neighbouring instances of the output artefacts remain similar. How much the Perlin noise affects an area of the screen at any one time is determined by the eye tracking data, as explained in the next section. There is however a very weak ‘gravitational pull’ on the parameter values towards their average. This ensures that the output remains continuous even after long cumulative exposure to the Perlin noise, preventing the generated outputs from fragmenting and separating from each other, as that would yielded visually obvious rifts on screen. When no gaze is detected, the system does a kind of random ‘drunken walk’ through the space of possible parameter values.

This is illustrated in fig. 4.6a, which plots how one parameter in *Keyebernates* changes over time when there is no gaze detected. The parameter value of each instance of the generative system is plotted as a very transparent black pixel, such that the more amounts of particular parameter value is present on screen, the darker the corresponding value in the plot. The parameter values are plotted on the x-axis, and time is on the y-axis, with time going downwards. When the system starts, all parameter values are the same, as can be seen at the top of the plot. As time progresses the variety of the outputs on screen increase in a smooth continual fashion, very much like a plume of smoke.

(a) (b)

Figure 4.6: Visualisation of the distribution of parameter values in *Keybernates* in (a), when no gaze is detected, and (b) with gaze interaction. Time is on the vertical axis, going downwards. Parameter value is on the horizontal axis. The darkness of the pixel represents the density of artefacts with parameters at this value.

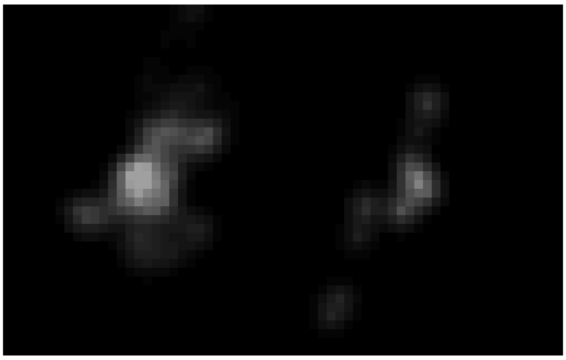


Figure 4.7: Heat map representing a rolling average of gaze fixation in recent past

4.2.4 Integration

To determine which parameter values are being attended to more than others, the stream of eye-tracking data is processed to calculate a continually moving 'heat-map' representing the fixation areas in the recent past (approximately 1 second, see fig. 4.7). This process acts like a de-noising, high-pass filter, smoothing out rapid saccadic eye movements and emphasising fixations. As studies mentioned in the section 4.1.1 suggest, gaze fixations and duration are supposedly the best indicator of preference, in this way the heat-map helps mitigate the Midas touch. This heat-map is mapped to the parameter planes to identify the parameter values of the artefacts most attended to.

Each frame, the further an area is from these parameter values, the more it is susceptible to the differentiating forces of the Perlin noise. This has the effect of slowing and eventually freezing the parameters of artefacts similar to the ones most looked at. Due to the weak 'gravitational pull', or bias, in the change of parameter values in the direction of the average value for each parameter plane, the range of parameter values represented on screen at any one time is constrained.

If gaze is fixated at just one point, gradually the range of parameter values represented decreases until all the artefacts on screen are identical. Gaze fixations thus engender a kind of 'zooming' in the abstract planes of parameter values, enabling a small area of parameter space to be explored in greater detail. A similar zooming process is employed in the text input system Dasher(Tuisku et al., 2008), when using gaze for character selection. By gazing towards the desired character, the interface zooms and the area around the fixation grows, making it easier to select the intended character. In *Keyebernetes* gaze also drives a zoom, but a zoom in *parameter space*.

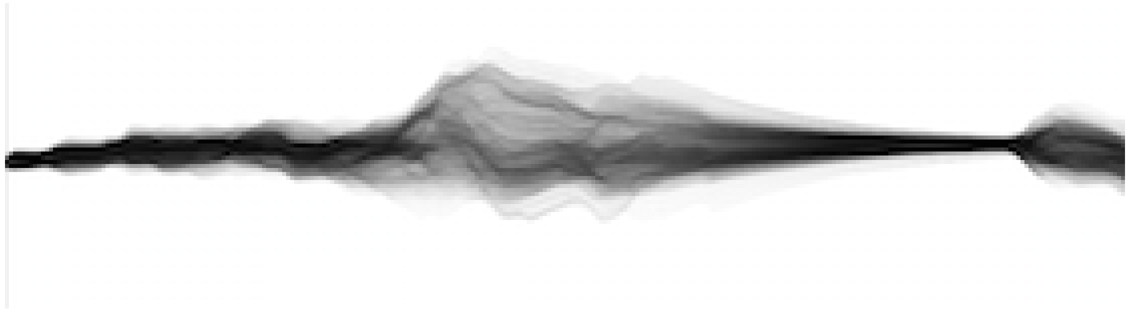


Figure 4.8: ‘Pinch’ of parameter values resulting from a steady gaze fixation.

4.2.5 System Dynamics

As *Keyebernates* is observed by a viewer, the gaze points counterbalance the randomising noise processes. Fig. 4.6b shows how a typical distribution of parameter values changes over time with the integration force of gaze. Like in fig. 4.6a, time is on the vertical, and the distribution of parameter values is on the horizontal axis. Note how the parameters sometimes seem to be ‘pinched’ and the variety in parameter value decreases (more detail can be seen in fig. 4.8). At these moments the gaze is fixed on one parameter value; a kind of ‘zoom’ in parameter space, and the more pinched the values, the more uniform the rendered output on screen.

A typical navigation through the parameter space is plotted in fig. 4.9. In these plots time is colour, black the beginning of the run and light green being the end of a 5 minute run with the overlapping circles generative output described in 4.1.1. Each axis corresponds to one of the parameter values. As can be seen, in this instance over the five minute navigation, there were increases in P_1 (*circle size*) and P_3 (*offset in y*), and a decrease in P_2 (*offset in x*). This plot illustrates how the navigation is akin to an exploratory ‘drunken walk’ through parameter space, however not completely random but shepherded by gaze fixations.

To further illustrate the dynamics of the system, consider fig. 4.10 which shows the navigation of one parameter over time. Here the red and green lines show the bounds of the parameter on screen (in this case P_1, circle size), the blue line shows the mean parameter value on screen. The light blue dots correspond to parameter values behind the gaze points. Note how these values are for the most part above the mean, and how this then pulls the parameter values upwards.

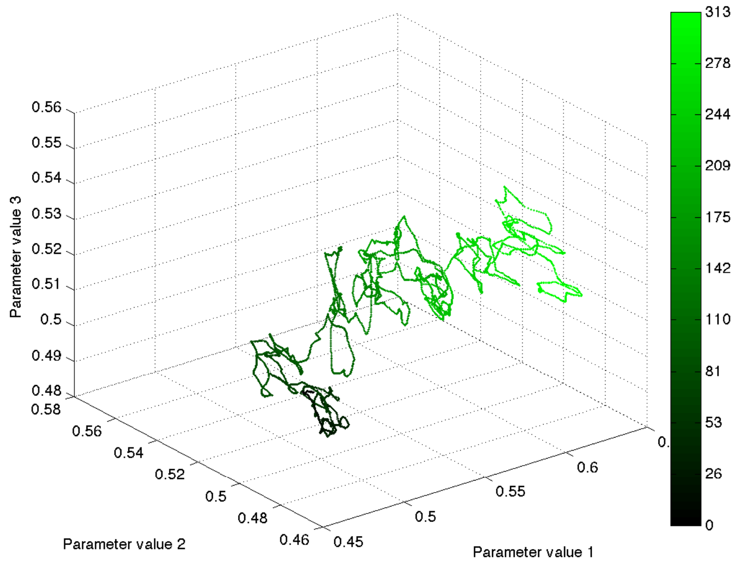


Figure 4.9: Mean value of parameter planes during 5 minute sample run. The colour represents time, with black being the beginning and green the end of the run.

Figure 4.10: Progression of P_1 over a 5 minute run of *Keyebernetes*. The blue line represents the mean value of the parameter on screen, the green and red lines show the maximum and minimum value on screen at any on time. The light blue dots represent the parameter values gazed at. Note how the gaze is usually fixated above the mean, pulling the parameter values upwards.

4.2.6 Meta-parameterisation - Balancing Keyebernates

For *Keyebernates* to work as intended, the counterbalancing forces of the Perlin noise and gaze need to be carefully calibrated. If any one of these forces is too strong or too weak, then the system as a whole is unable to engender the navigation through solution space. For instance, if the Perlin noise affects the parameters too strongly, then the pull of the gaze has little effect and the navigation through space is essentially random and chaotic. Conversely if the magnetic pull of the gaze is too strong, the parameters all converge, and then the system gets stuck in a state of stasis, unable to change.

To achieve this, an iterative process of manual calibration - of *meta-parameterisation* - was carried out to ensure that when looked upon, the dynamics of the system would reside in this space between stasis and noise. The meta-parameters that go into calibrating *Keyebernates* are numerous. These include everything from the dozens of parameters that defined the Perlin noise, how smoothly to interpolate gaze points, how strong each gaze points affects the 'heat-map' which, in turn affects the parameter-planes' tendency to be affected by the noise, or the strength of the overall gravitational force that keeps the parameters planes from fragmenting.

Further finding the right meta-parameters is very specific to the exact technical substrate; the performance and speed of the computers involved (and the network between them), and the rate at which the gaze-tracking system is able to transmit gaze points (approximately 60hz) for instance all have an effect. A large amount of trial and error, in an iterative process was needed in to find these numerous meta-parameters.

Keyebernates is a generative system itself. Even though an aim of this system is precisely to enable parameter discovery in a smooth, continuous fashion, calibrating *Keyebernates* is itself subject to a discontinuous, steps of parameter selection, and parameter evaluation. As such the final meta-parameters selected were chosen because they seemed to work well, rather than due to any objectively measurable criteria.

4.2.7 Implementation

The eye tracking system used is Face Lab 5 by Seeing Machines[11], a non-contact optical tracker. A binocular set of 60Hz infra-red cameras placed below the screen detect the reflection of an infrared light in the eyes of the subject, and sends in real-time the screen coordinates of the subject's gaze points over a network to another computer running to the main *Keyebernates*

[11] www.seeingmachines.com

application, that in turn interprets the gaze data and renders the visual output. The eye tracker offers an error of less than 0.5°, which results in 5mm diversion from the gaze point with the user looking at the screen from a distance of 50cm. The screen had resolution 1280x1024 pixels. *Keyebernates* was developed with the OpenFrameworks middle-ware layer which is built on top of C++.

4.3 Study

As mentioned earlier in section 4.1.1, gaze is understood to be "both simultaneously bottom-up, stimulus-driven as well as top-down, goal-oriented"(Duchowski, 2002). The distinction between volitional and non-volitional gaze informed the design of the study so that both scenarios could be observed. The first part of the study had the participants passively observe the happenings on screen, in the second part of the study participants were asked to look for a particular pattern. The participants were not told that their gaze affects the output until after the end of the experiment, and rather were lightly deceived and told that they would be watching a video.

The output for this study was a very simple parametric design consisting simply of overlapping circles described in section 4.2.1.

For all tasks the system started in the 'middle' of the state-space (all parameter planes at 0.5), with the pattern shown in fig. 4.16a: medium-sized circles with no offset. There was a total of 6 tasks, 3 in the passive observation part, and 3 in the search part. The order of the tasks were randomly shuffled for each part. Each task lasted 5 minutes. However some participants reported getting fatigued, so this meant that only 4 out of the 6 participants completed all tasks.

The system was set to pause when no gaze was detected, allowing the participants to rest their eyes should they feel the need, and the system would resume where they left off when they returned their gaze to screen. In the search tasks, the participants were shown the target pattern printed on paper. By having the system pause itself, there would be no adverse effects when they look to the paper for reference.

4.3.1 Participants

The study was carried out with 6 subjects. The subjects of this study were three males and three females, all postgraduate students between 26 and 35 with normal uncorrected vision. For each subject, a calibration of the eye tracker was carried out, this is a simple process where by the

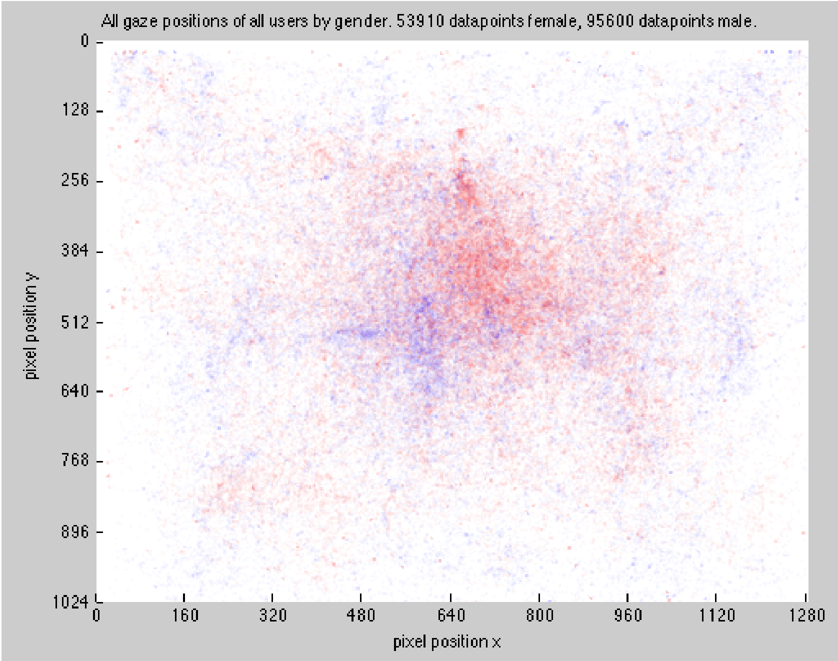


Figure 4.11: Plot of all the gaze points from all subjects across all experiments. Red points are from the male subjects, blue points are from the female subjects

viewer follows a dot on screen.

4.3.2 General Gaze Patterns

Fig. 4.11 shows all the gaze points of all users across all experiments. There are more points from male users as more of them carried out all parts of the experiment. The red points are from the male viewers, and the blue points are from the female subjects. As can be seen, the middle areas of the screen had the most attention, with areas leading up to the upper corners also having more visual attention, leading to a sort of butterfly shape. This is not unexpected, as the corners, with the both vertical and horizontal edges, can be viewed as having more 'information'. Areas that had the least gaze points were along the middle of the edges of the screen, particularly the lower-middle left edge.

There is a notable gender difference; the males tended to look at the right hand side of the screen and while the females looked more evenly. Further work is needed to establish if this gender difference holds more generally, as the number of subjects in this study is small. Some literature has shown that there is a left-right bias preference, with the right-hand side being preferred in right-handers (e.g. (Beaumont, n.d.)). All subjects in the study were right-handed. With regards to the gender bias, a study by Drake(Drake, 1987) testing left-right bias on aesthetic judgment of photographs was able to detect this bias in males but not in females. Perhaps this indicates that there could be other uses for *Keyebernates* as an experimental stimulus in perception research.

4.3.3 Experiment 1: Passive Observation

In this experiment, subjects were asked to simply observe a 'video' on screen. They were not told that their gaze affected the output. This was repeated two times for each subject, with *Keyebernates* setup with different meta-parameters for each run. The difference between the setup for each run was the strength of the integrating gaze force.

In *case 1* the integrating force was 'weak' relative to the effect of the noise, and in *case 2* the integrating force was 'strong'. This difference in integration force strength changes the amount of time it takes to make all the parameter planes uniform, by fixating the gaze on one point of the screen. This is the duration that it takes to 'pinch' the values in parameter planes as shown in fig. 4.8.

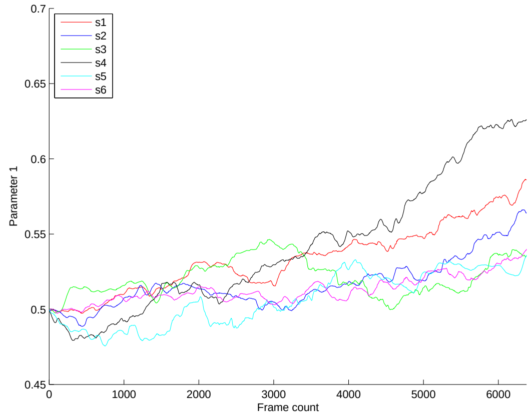
In the 'weak' case, *case 1*, the time it takes for the parameter planes to become uniform from a maximum heterogeneity is approximately 5 seconds. For the strong case, *case 2*, the parameters would be pinched faster, after approximately 2.5 seconds.

In addition to getting some insights into the dynamics of the system in use, an aim was to see if multiple subjects end up taking similar routes through state space.

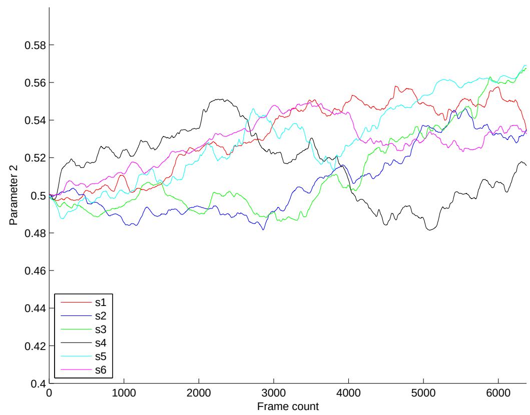
4.3.4 Experiment 1: Results

In figure fig. 4.12 plots for each subject's navigation in the passive tasks for the 'weak' gaze case are shown, and in fig. 4.13 for the 'strong' gaze case. Each colour is a subject, and each plot corresponds to one of the three parameters. Time is on the x-axis, and covers five minutes, and the parameter value is on the y-axis. As mentioned, the offset values represent an onscreen shift in x or y for the circle, but this shift would be inverted each row/column. As such, reflecting an offset value across an axis at parameter 0.5 would yield the same perceptual output, only shifted by one row or column.

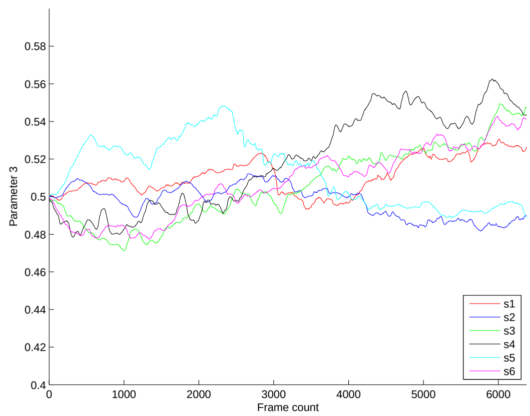
As can be seen, the plots in fig. 4.13 are 'noisier' than the plots of fig. 4.12. This due to the stronger effect of the gaze, pulling the parameter values faster. The clearest trend is that in these five minute navigations through the parameter-space, the subjects pulled the value of P_1 upwards, in both *case*1 and *case*2. This indicates that the subjects' gaze is more attracted to the larger circles. However with regards to the offset parameters, P_2 and P_3, there was only moderate



(a)

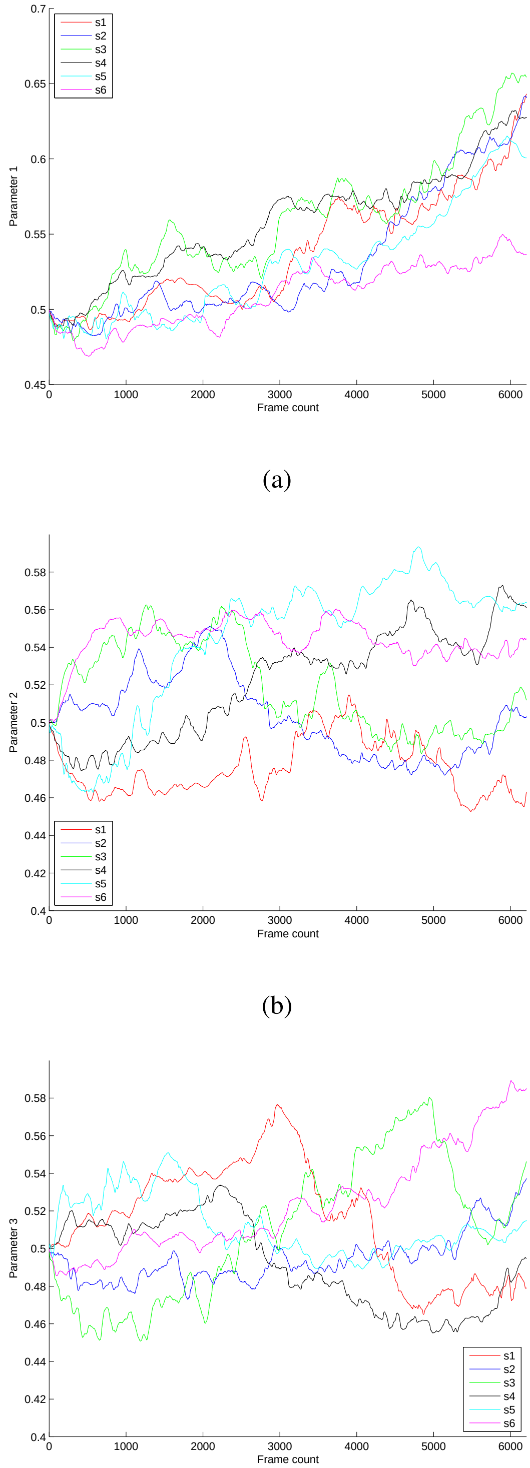


(b)



(c)

Figure 4.12: Plots of the mean parameter values for all subjects in *experiment*1, *case*1; passive observation with 'weak' gaze integrating force. (a) is P_1, circle size. (b) is P_2, offset in x. (c) is P_3, offset in y.



(a)

(b)

(c)

Figure 4.13: Plots of the mean parameter values for all subjects in *experiment*1, *case*2; passive observation with 'strong' gaze integrating force. (a) is P_1, circle size. (b) is P_2, offset in x. (c) is P_3, offset in y.

shifts away from the 0.5 centre. The furthest these parameter ever got from was a distance 0.06 in *case*1, and about 0.09 in *case*2, significantly less than the distances travel with P_1, up to 0.13 in *case*1 and 0.16 in *case*2.

In fig. 4.14 and fig. 4.15 the generated output over the 5 minute run for each subject is displayed. As can be seen, as the circles increase in size, there is more overlap between them. This yields more complex and intricate patterns. This suggests that an increase in the phenomenal complexity results in a more 'eye-catching' the result. Glancing at the bottom of these plots indicates the final result of the navigation after the five minutes.

Overall the distance traveled in parameter space was not great. This is because the design of *Keyebernates* relies on visual changes in the output to occur gradually, making it best suited for a slow, but detailed exploration of a small sub-set of the parameter space. To do an exhaustive exploration through this modest three-parameter design would take a prohibitively long time with this interface. Possible approach to dealing with this restriction are outlined in section 4.4.

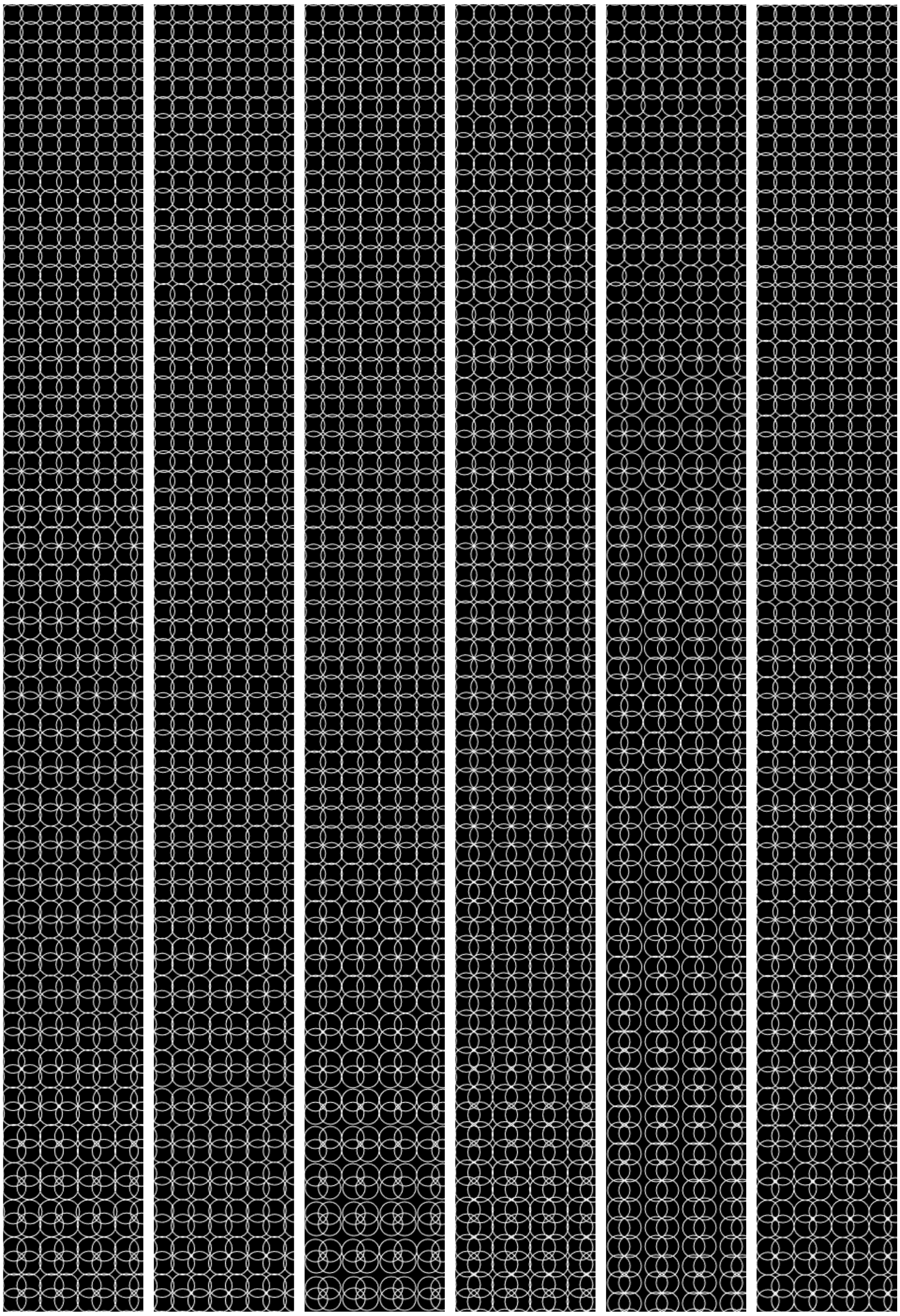


Figure 4.14: The mean generated output for *experiment*1, *case*1, over the five minutes for each subject. Time is on the vertical, with the beginning at the top. Each column is the output of a subject, beginning with *s*1 to *s*6.

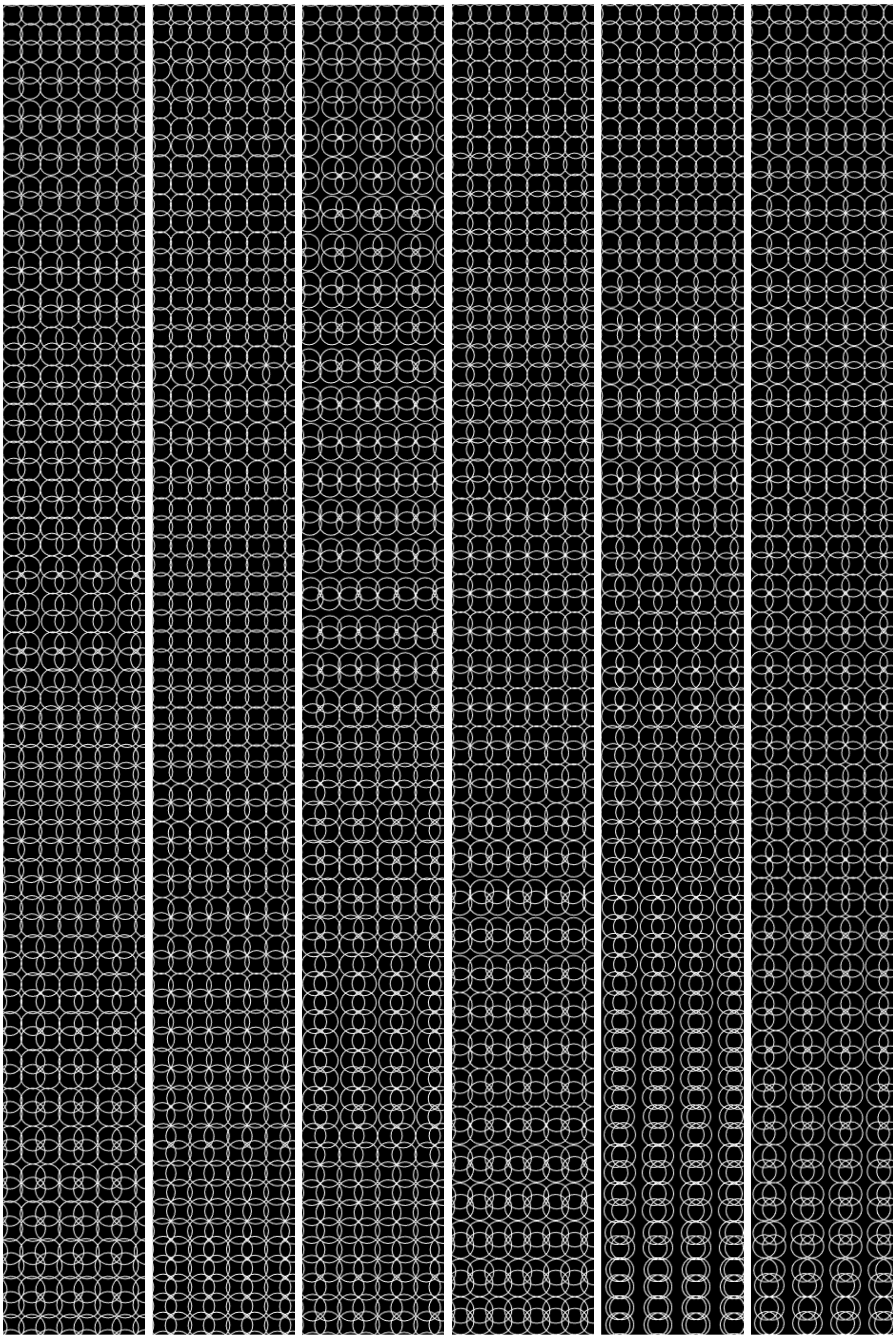
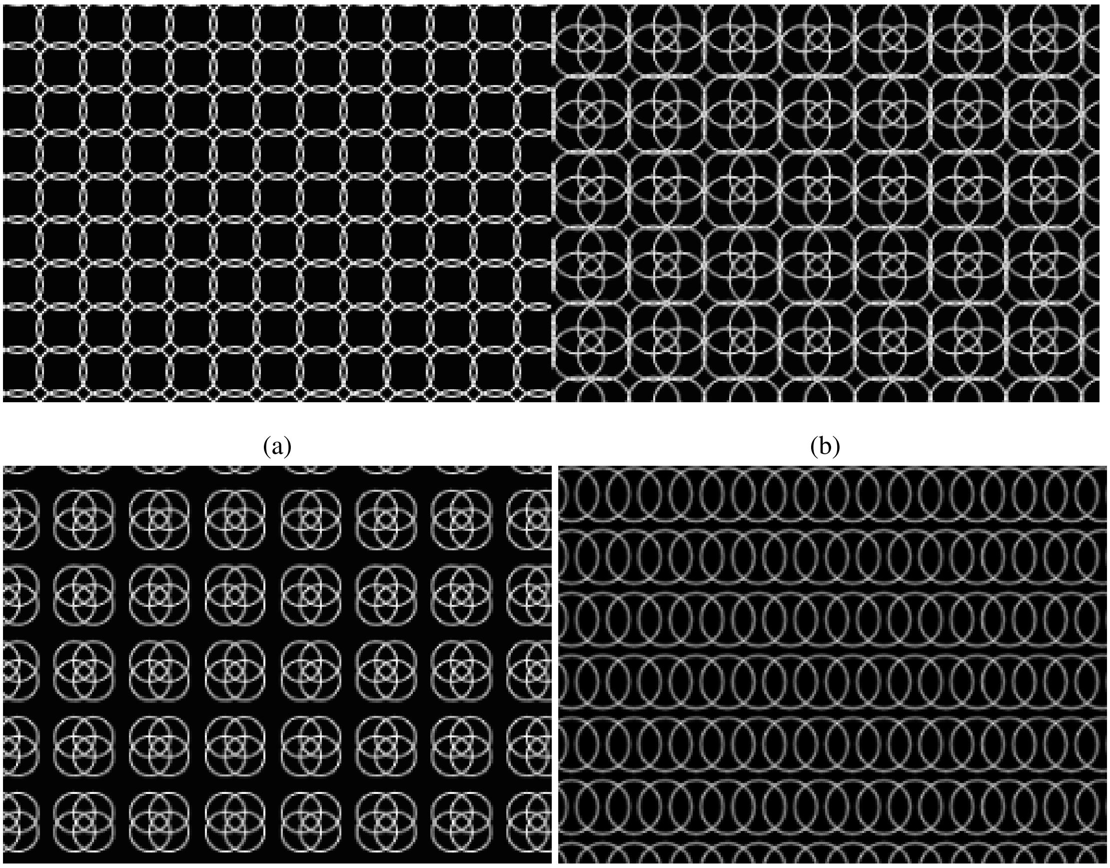


Figure 4.15: The mean generated output for *experiment*1, *case*2, over the five minutes for each subject. Time is on the vertical, with the beginning at the top. Each column is the output of a subject, beginning with $s1$ to $s6$.



(a) (b) (c) (d)

Figure 4.16: Search targets for *experiment*2. (a)The starting position at the middle of the state space (P_1=0.5, P_2=0.5, P_3=0.5). (b) The 'near' search target (P_1=0.64, P_2=0.55, P_3=0.44). (c) The 'middle' search target (P_1=0.44, P_2=0.4, P_3=0.59). (d) The 'far' search target (P_1=0.83, P_2=0.5, P_3=0.33).

4.3.5 Experiment 2: Search Task

In this part of the study, subjects were asked to search for specific patterns showed to them on a printed piece of paper. It was explained to them that in the 'video' they are observing, the target might or might not appear, but to attempt to spot it as soon as they could.

There were three search targets (see fig. 4.16). Each target is a point in parameter-space, and the Euclidean distance between these and the starting point in the middle of the parameter-space varied from 'short', to 'medium', to 'far' (with euclidean distances 0.16, 0.2, 0.37 respectively). The subjects were told that the pattern they are looking for 'might or might not' appear, but if it does to try to spot it as soon as possible.

Like in *experiment 1*, the system was setup to automatically pause if the viewer averted their gaze from the screen, allowing the subjects to both rest their eyes, or to refer to the print out of the search target as often as they pleased. Again they were not told that their gaze affects the

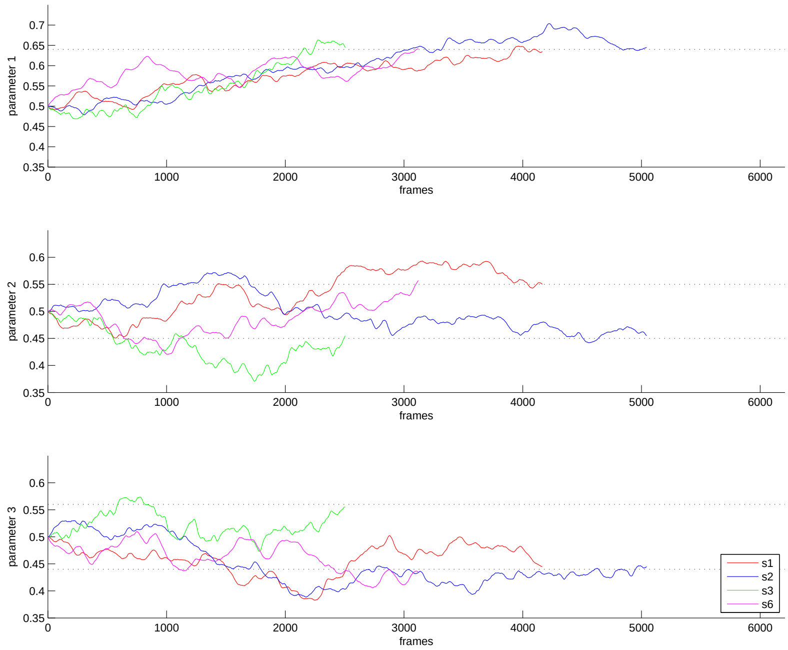


Figure 4.17: Parameter values for the search task to the 'near' target, for P_1, P_2 and P_3 respectively. Time is on the x-axis, representing five minutes, and the value of the parameter is on the y-axis. The dotted lines represent the target's parameter values. Each coloured line represents the path taken by a subject.

output. Each search had a time limit of five minutes. If the target parameters appeared anywhere on screen, the experiment would stop automatically. The order in which the tasks were carried out was shuffled for each subject.

4.3.6 Experiment 2: Results

Fig. 4.17, fig. 4.18 and fig. 4.19 show the trajectories of the parameter values for the 'near', 'middle' and 'far' targets respectively. The dotted black lines show the target parameter values. Note that for P_2 and P_3 there are two 'correct' answers, due to equivalence of the parameters when reflected across the 0.5 axis.

As can be seen in fig. 4.17, all four of the subjects who under took this part of the study ended up steering the system to the 'short' search target. The quickest, $s3$, did so in just over two minutes, while the slowest $s4$, took twice as long. As can be seen the subjects navigated quite differing paths in parameter space before reaching their target.

Fig. 4.18 shows how one subject, $s3$, reached the 'middle' target with another $s1$ getting tantalisingly close. Others however seemed to not get so close, with $s2$ having gone in the 'wrong

Figure 4.18: Parameter values for the search task to the 'middle' target, for P_1, P_2 and P_3 respectively. Time is on the x-axis, representing five minutes, and the value of the parameter is on the y-axis. The dotted lines represent the target's parameter values. Each coloured line represents the path taken by a subject.

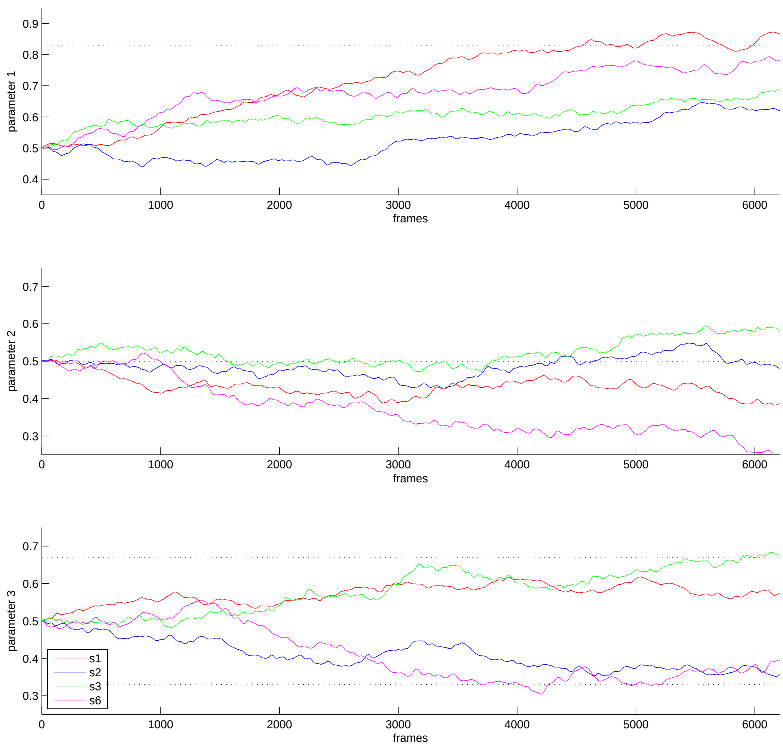


Figure 4.19: Parameter values for the search task to the 'far' target, for P_1, P_2 and P_3 respectively. Time is on the x-axis, representing five minutes, and the value of the parameter is on the y-axis. The dotted lines represent the target's parameter values. Each coloured line represents the path taken by a subject.

direction' with respect to the circle size parameter right from the outset.

None of the subjects reached the 'far' target in the five minutes allocated, as can be seen in fig. 4.19. None of the subjects could be said to have gotten particularly close to reaching the target, and at the end of the five minutes, some were clearly on a trajectory leading even further away.

These results seem to indicate that gaze can guide *Keyebernetes* when the individuals are looking for a particular pattern, however only when the parametric distance is small. Individual difference in the performance of this task indicate that gaze patterns were quite different for each of the subjects.

Because the parametric design used in this experiment consisted of overlapping entities there is no guarantee that visual similarity coincides with parametric proximity. Consequently there are cases where the output at a certain point in parameter space is visually similar to one far away. So searching based on visual similarity could lead to a 'dead end' in the navigation, and the number of such 'dead ends' increases the greater the parametric distance to the search target. However over short parametric distances, visual similarity coincides with parametric proximity, and the results of this experiment indicate that at these scales it easy to lead system to a target output just by 'looking for it'.

4.3.7 User Feedback

Fatigue as an issue has been noted in numerous eye-tracking studies. Here similarly fatigue was a problem for some of the participants, with only four of the 6 participants feeling comfortable carrying out all parts of the study.

In this case, it is likely that the choice of parametric output contributed further to the fatigue, as the output in the tasks was variously reported as being ‘trippy’ and ‘intense’ and as consisting of “too much input”. Subjects expressed that some of the fatigue could have come from the cognitive load of trying to work out what the animation was doing. One subject reported for instance that “I knew it was changing, but I couldn’t actually see it change, I was trying to work it out”. Some subjects reported perceiving a shifting ‘blur’ in the peripheral vision (as the artefacts changed in parameter value), and that this contributed to their confusion. Most subjects found that the search tasks were less tiring than the passive observation.

After the experiment, the subjects were debriefed, and only then was it revealed how their gaze was controlling the animation. They were then told how to use their gaze to consciously make the patterns at one area of the screen to spread to the rest, which all subjects could do easily.

4.4 Discussion and Further Work

In summary, the following observations were made over the course of the experiments:

- *Observation 1*: Overall gaze patterns favoured the centre of the screen, followed by the upper corners, with a possible gender bias of right-hand preference in males.

- *Observation 2*: Non-volitional, passive observation gaze is attracted to more intricate patterns, as a consequence of greater circle intersection due to the larger circle size.

- *Observation 3*: Volitional, searching gaze can guide the system across small parametric distances.

- *Observation 4*: Fatigue is a significant issue.

4.4.1 Success

Could *Keyebernates* be considered a ‘success’? Was it successful in mapping out the parameter-space of the generative process with respect to cross-user, objective aesthetic desirability? A metric for success was defined in section 1.2.1. It consists of two criteria: $C1$ - *correlation* *between parameter choices and aesthetic value of the output*, and $C2$ - *sufficient volume of data points and significant trends*.

C1

With regards to *observation 2*, care must be taken before attributing greater aesthetic value to the patterns with larger circles, or claiming that passive observation of *Keyebernates* necessarily engenders a navigation towards parameters yielding objects of greater aesthetic worth. Certainly, when comparing the starting patterns of *experiment 1* with the finishing patterns (as can be seen by looking at the top and bottom of fig. 4.14 and fig. 4.15), it is easy to see that the finishing patterns are certainly more intricate. As the circles increase in size, they overlap more, yielding more complex patterns.

Perhaps it is true that gaze is attracted to the more aesthetically pleasing objects, as some of the research surveyed in section 4.1.1 suggests, but the results of *experiment 1* are unable to verify this. *Keyebernates* only displays a small range of the generative artefacts at any one time, and never are instances of the artefact that are far away in parameter space compared against each other on screen.

It is clear that for all subjects, passive gaze pulled the system towards larger circles and thus more intricate patterns. However that does not mean that patterns with larger circles are more beautiful, only that they are, generally, more 'eye-catching' than patterns in neighbouring areas of parameter space. This study was carried out in a lab setting; the environmental conditions were controlled, and the screen, its settings, as well as the configuration of the eye-tracking system was identical for all participants. There is in a sense the inverse problem to the issues of medium-specificity in the other experiments. In the other studies, data collection through crowdsourcing necessitates a lack of control over the presentation conditions for each individual. Whereas in this study, the presentation conditions are the same for each individual.

It is difficult to know if the results of this study would translate to different viewing conditions. Would the gaze still be drawn to larger and larger circles if the display were larger? What about if it were brighter? Or darker? What if the colours were inverted, and the background was white and the circles black? And indeed what if the work was presented in an entirely different context, and rather than being in a lab setting, was presented in an art gallery? It is not possible to answer these questions currently, however they must be kept in mind when considering the results gathered: the results of these analysis, can only really be considered valid for the particular setup of this lab-based experiment.

It has not been possible to attribute this attraction to the larger circles to aesthetic preference, as such it cannot be said that this experiment could meet the criteria $C1$.

C2

The second criteria – *sufficient volume of data points and significant trends*– was also not met. There were two main difficulties: the issue of fatigue made it such that participants were not able to interface with *Keyebernates* for long durations, and the inherent slowness of the navigation through parameter space that *Keyebernates* engenders. These two factors made it such that only a small portion of the parameter space could be explored by the participants.

A fundamental limit of *Keyebernates* is that it only allows one area of parameter space could be explored at a time. If this area is near a local maximum of gaze attraction, then it becomes difficult to escape this local maximum. The gathered data did not indicate that such a peak had necessarily been reached, as the navigations did not seem to have stabilised. However the exploration is very much local to one area of parameter space, and to be able to get a sense of the geography of the fitness landscape of the generative process, there would need to be a mechanism for exploring disparate areas of parameter space.

One way of achieving this is to use *Keyebernates* to periodically jump to a completely different area of the parameter space, rather than have the participants start in one place and just navigate the area nearby. Blinks can be detected in real time with eye-tracking technology. For instance every 20th blink, *Keyebernates* could jump to random point in parameter space, and let the navigation begin from there. This would allow for disparate parts of the parameter space to be mapped out. The issues of slowness of navigation and fatigue would still need to be overcome, however this could make it possible to identify families within the parameter space that are most attractive to gaze.

However *Keyebernates* in its current form, as it failed to meet both $C1$ and $C2$, was not successful in uncovering the underlying landscape linking objective aesthetic desirability to the parameter-space of the generative process.

4.4.2 Future Work

Observation 3 shows that there is some potential for *Keyebernates* as a design interface, however a number of issues must be overcome before it can become a *practical* interface, a subject of further work.

The results of *experiment 2* clearly show that the gaze directed navigation is only possible across small parametric distances. As *Keyebernates* relies on changes in the peripheral vision to be slow and continuous, the speed of navigation in parameter space is necessarily very limited. One of the issues is that only a small amount of parameter-space can be shown on screen at any one time. This problem gets worst the more parameters the generative process has; as there is a limited, and constant, amount of screen-real estate, for each additional parameter added to the generative process, the span of values for each parameter renderable on screen decreases as the number of dimensions to be navigated increases.

The design of *Keyebernates* is such that it is more suitable to slow detailed explorations of a small area of state-space. Further, the issue with the afore mentioned disconnect between parametric similarity and phenomenal similarity makes it hard to anticipate in what 'direction' in parameter space one is to navigate to reach a particular phenomenal result.

Mitigating Fatigue

Keyebernates is gaze-contingent; it behaves and appears differently with respect to how it is gazed upon. However there is nothing gaze-contingent in nature; nowhere is there a situation whereby something moves only when not directly pointed at with gaze. It is in this sense a highly *unnatural* situation. It is then perhaps unsurprising, as noted in *observation 4*, that interfacing with *Keyebernates* is fatiguing; the viewer's visual system is continually trying to build a coherent representation of the scene, while the scene is shifting constantly, but only in the peripheral vision.

It may be that solving this issue would involve changing the dynamics of *Keyebernates*. In particular, if the changes in the peripheral vision were such that they were more subtle, perhaps interfacing with it would be less fatiguing. Further work could explore this by changing the rate of changes in *Keyebernates*. It is possible that making change happen an order of magnitude slower could mitigate the fatigue issue, as the activity in the peripheral vision may be harder to detect. This would conflict with making the interface practical in any sense, as the navigation through parameter space would be slower still. However, if *Keyebernates* is only to be considered an artistic object, as discussed below in section 4.4.3, than practicality may not necessarily be a concern.

Another possibility is that working with a higher resolution displays might mitigate fatigue. As the sizes of the circles change, they naturally quantise to the nearest pixel. These quantisation steps are noticeable sudden changes, but have been minimised through blurring of the output, as shown in figure 4.4. A higher resolution display – such as modern 4K ultra-high definition screen – would allow for smoother, more continuous and even less noticeable changes.

If the aim is to make an interface to assist designers in parameter discovery, these experiments, and the literature, show clearly that eye tracking as the sole input modality is not practical. Fatigue could be mitigated if eye-tracking was one of a number of ways of interacting with the system . For instance, a physical haptic controller, such as motorised sliders[12], could allow the designer to quickly navigate to any point in the state space with their hands. By assigning each slider to a parameter value, it could either be moved by the designer or could be made to move by themselves to reflect the state of the system. The designer could for instance find an approximate design manually, and then put one or more of the parameters under gaze control to explore a detailed area of the state-space. Designers do not have clear search targets in mind when working with parametric designs (then they would have already finished their design!), but if they 'know it when they see it', perhaps this framework could assist in a design context.

In response to *observation 1*, an intriguing possibility that could help further improve the performance of the system involves leveraging the overall gaze patterns outlined in section 4.3.2. These patterns could be used to construct a 'map' to give greater or less importance to fixations at different areas of the screen. For instance if a fixation occurs near the lower/middle left edge of the screen, it could be given greater importance than a fixation at the centre of the screen.

Re-mapping Parameter Space

Another possible approach to improve the performance of *Keyebernates* is to provide an intermediate mapping layer between the parameter planes and the generative artefacts. Currently sets of parameters that are far apart in parameter space can yield outputs that are similar (as an extreme case consider the reflections in the offset in x and offset in y across the 0.5 axis), conversely some small differences in parameters values can yield perceptually large differences. There is a non-linear mapping from the parameters to the qualities of the work; a viewer does not experience

[12] As commonly seen on modern mixing desks, where the sliders can be made to move in response to commands from a computer.

'offset in y-ness', but rather might experience 'floweriness' or 'row/column-ness' for example.

If one could identify properties of the *conceptual spaces*(Gärdenfors, 2004b) of the generative system at hand, then it could be used to create an alternate parameter mapping. The outputs of the generative processes used in these experiments, as can be seen from the examples, can sometimes look a bit like flowers of sort, that could be labeled by an observer as 'flowery'. However there is nothing in the definition of this generative system *per se* that implies floweriness, that an output is 'flowery' is entirely a property of the perceiver's interpretation of this output.

If the parameter mapping could be re-aligned to match the relevant quality dimensions, then it would be possible to place all the 'flowery' parameter values in adjacent areas of this re-mapped parameter space (defining what Gärdenfors calls a 'natural concept'). The mapping between the 'floweriness' quality dimension and the parameter space is non-linear and discontinuous; roughly speaking flowery outputs will be the outputs with a circles size within a certain bounds (the circles need to overlap, but not too much), with the *offset in x* and *offset in y* being of roughly equal distance from 0.5.

If such a mapping were implemented, this would make it such that similarity *would* coincide with parametric proximity, and then searching for a particular type of design could potentially be much easier and faster; navigations would be less likely to get stuck in 'dead-ends' and it would be easier to determine in what direction in parameter space one is to travel.

Such quality dimensions could conceivably be applied to any generative process. An awareness of the relevant quality dimensions is of value to the designers and composers who work with any particular generative system, and is naturally built up as the designer/composer becomes more familiar with the generative process or the formalisms within which they operate. The *Melody Triangle* was designed to explicitly provide a linear mapping between input parameters and the 'quality dimension' of predictability. The mating algorithm of the *EvoColour* interactive evolutionary system project described in the next chapter operates with respect to the quality dimensions of the perceptual similarity of colours.

However it is unclear how one could build such a 'quality dimension mapping' for *Keyebernates*. For the *Melody Triangle* it was based on extensive knowledge and theory of the process at hand. For *EvoColour*, the quality dimensions of colour leveraged in its algorithms were identified through the history and evolution of colour theory over centuries of research.

But for any given generative process, how is one to identify the quality dimensions? Gärdenfors makes it clear that quality dimensions and conceptual spaces are based one phenomenal *similarity*(Gärdenfors, 2004b). One could conjecture that if a large number of similarity measurements were made on randomly selected outputs of a generative system, these measurements could then be used to build this re-mapping of parameters. This proposition is the subject of further work.

4.4.3 *Keyebernates* as a Work of Art

Clearly the value of *Keyebernates* as a practical interface is not established, certainly in its current configuration. Further as an object of study in empirical aesthetics, it is not practical for reliably extracting parameter values with respect to the aesthetic value of the output.

From a cybernetic point of view, *Keyebernates* is a *first-order*, reactive system(Dubberly, Pangaro, & Haque, 2009). It is in some sense a-kin to a thermostat, trying to adjust the 'temperature' (i.e. output) by sensing the 'environment' (the 'preferences' of the viewer, as interpreted by the gaze).

A thermostat operates along a one-dimensional parameter space, mapped exactly to one quality-dimension: temperature. However *Keyebernates* operates along a multi-dimensional parameter space, of as high a dimension as there are parameters in its output generative process. Anthropomorphising *Keyebernates*, it could be said that it is constantly (if naively) 'aiming to please' as much as possible.

There is a certain poetic weight to the idea of an image that 'aims to please', of an image that changes, but without it being perceived to do so; a work that never ends. One must await the days of more sophisticated, unencumbered and subtle gaze tracking technology, for this work to be realisable in an art gallery context[13], but one could imagine it there, without it being known what it will do, changing constantly, slowly and invisibly.

Attribution and Creative Responsibility

Why might it be desirable to automate, or rather, make subconscious, parameter search with gaze in this way?

Decisions and choices in art are often difficult, but most crucially, are usually taken by *someone*, even if indirectly through algorithms or generative processes. A key aspect of *Keyebernates* is that the viewer should not know that their gaze affects the output, because otherwise it then

[13] *Eye-Machine-Bot*, mentioned earlier, is the only instance of eye-tracking in an art gallery context that could be found. However it was far from unencumbered, with visitors needing to place their heads in brace and carrying out a calibration task before the work could begin.

becomes just another interface for control (and an impractical one too). When a viewer is observing *Keyebernates*, where then does the decisions over which parameters are to be reinforced come from? Who is responsible for the choices made and the consequently generated output?

The meta-parametrisation of *Keyebernates*, described in section 4.2.6, balancing the *integrating* and *differentiating* forces, is akin to the 'tuning' processes seen in 'ecosystemic' installations, such as in the work of Agostino Di Scipio, Gordon Pask or Alvin Lucier (surveyed in section 2.2.4). Here the composer/designer carefully adjusts the elements that affect the environment (actuators), and those that sense the environment (sensors), such that the combined behaviour of the system elements in the environment would yield apparently emergent, novel behaviour. These kinds of systems aim to be in the 'middle of the triangle', not too predictable but not too chaotic, to reside 'Entre le crystal et la fumée'(Atlan, 1979), the realm of living entities.

Despite being coupled to a specific subject, *Keyebernates* does not elicit their volition, rather their preconscious, objective gaze drives the decisions. A particular path taken through parameter space, and the generated outputs belong to each individual viewer. Paradoxically however, they cannot be held responsible for the decisions made. Although the present author is responsible for setting up the situation for this navigation to take place, once it begins, there is only the interaction between the noise processes and the perceivers' gaze that directs the navigation through parameter space. To whom is one then to attribute *Keyebernates* and the outputs it generates? As a whole, as this amorphous process where viewers unwittingly are coupled to an emergent exploration of parameter-space, *Keyebernates* as a 'work of art' in its own right, is best attributed to the present author. Just as in Di Scipio's *Audible Ecosystemic Interface*(Di Scipio, 2003), where 'the sound is the interface', in *Keyebernates* the generated output on screen – its light and patterns – is the interface. Just like a living organism, *Keyebernates* both seems to have its own will, and reacts to the environment. And like a living organism, it may do what is expected, or it may not.

In this chapter, *Keyebernates*, an experimental system where gaze would influence the evolution of parametric designs was presented. Studies were carried out to explore its behaviour under both passive observation, and in search based tasks. It was found that passive observers would pull the system towards images of greater phenomenal complexity. However it was not deemed to be 'success' at uncovering patterns of objective aesthetic preferences with respect to the parameters of its generative visual processes. It was shown that volitional-gaze could be used to navigate the parameter-space of generative designs across modest parametric distances. This suggests that such a system has potential as a possible interface for parameter discovery. However it is in its current form limited in the range and speed of parameter space explorations, and further the system was fatiguing to interface with. Possible steps that could be taken towards mitigating these issues were suggested as possible further work. Finally the value of *Keyebernates* as an aesthetic work in its own right was outlined.

In the next chapter, a crowd-sourced evolutionary system, *EvoColour* will be presented. Like *Keyebernates*, it conflates the role of consumer and designer; the viewers directing the course of the evolution, and the shape of the images. Further it also carries – within the design of its evolutionary algorithms – this kind of balancing of differentiating and integrating forces, where the volition of viewers make things more similar, but mutations make things different. And like *Keyebernates*, it is both an experiment in empirical aesthetics, as well as a work of art – or rather: a system that generates works of art.

Chapter 5

EvoColour

5.1 Introduction

This chapter concerns the third practical study of this research: *EvoColour*.

EvoColour is a crowdsourced, interactive evolutionary system where populations of simple images – consisting of a sequence of fifty concentric circles of up to three colours – 'evolve' in response to preference selections made by members of the public over the internet. Using a Darwinian algorithm, images that are more popular, as determined by selections of users on the web, are select for survival, while those that are least popular are 'killed off' (removed from the population). The popular images then sexually reproduce to fill the slots vacated by the killed-off images.

EvoColour falls under the umbrella of Evolutionary Music and Art (EMA), discussed in section 2.2.4 of the background chapter. However unlike many other EMA systems, which have a virtually unbounded range of outputs, here the space of possible outputs is purposefully small. Like *Keyebernetes*, *EvoColour* is not only a psychophysical experiment seeking to find relationships between measurable visual characteristics and objective aesthetic preferences, but is also a work of (metacreative) art in and of itself.

There are two fundamental motivations behind the creation of *EvoColour*. The first is that of scientific enquiry, the other is part of a personal artistic journey. Like the other practical studies of this thesis, *EvoColour* attempts to uncover the universal, objective elements of aesthetic preferences, with regards to a particular generative process.The evolutionary algorithm is designed to optimises across the collective judgments of the many, and in so doing removes the idiosyncrasies of tastes and the subjective elements of aesthetic judgments. What is left behind is a fitness landscape that reflects the universal elements of aesthetic taste (with regards to these particular kinds of images).

To automate and crowdsource parameter discovery in this way is useful not least because it can provide insights into mechanisms of human perception, but can also lead to numerous practical applications, particularly in design. The findings of this study can form the basis of design heuristics and inform the work of artists and designers. Additionally they could be incorporated into the construction of smart design tools that could provide informed feedback, or even stand in and autonomously make creative decisions that are likely to appeal to many.

The personal artistic motivations stem from a simple curiosity and desire to make objects of aesthetic worth. If the evolutions are able to elicit from the crowds common aesthetic judgments, what would such evolved images look like? Will they be striking and broadly appealing? Is it possible to find an 'archetypal image' that is of greater beauty than the rest?

EvoColour is accessed via a web site[1]. The home page provides a brief introduction to the system, links to further information and statistics, including a 'Hall of Fame' of users with the most selections, and a 'top 20' of the currently most popular images (fig. 5.1). Once the user has filled in a short registration form and begin the study, they are presented with series of pairs of images, as shown in fig. 5.2. The users simply select which of the two images they prefer by clicking on one of the images, or by using the arrow keys on the keyboard. Users could also participate via mobile devices, where they would identify there preference by tapping on the image. They could also indicate they have no preference by pressing the 'no preference' button. As soon as this is done, the user's selection is sent to the server running the evolution, and a new pair of images is sent back to the user's web-browser and rendered on screen. The users could do as few or as many selections as they choose. The accumulated selections of the users determine the image's fitness, which defines wether the image survives to pass on its characteristics to the next generation. If an image is selected, then its fitness score is incremented, and if the other image is selected, its fitness score is decremented.

There are alternatives to binary selections, for instance users could be asked to give a ranking of an images on a scale, or as in some other systems be presented with an array of images of

[1]http://www.evocolour.net

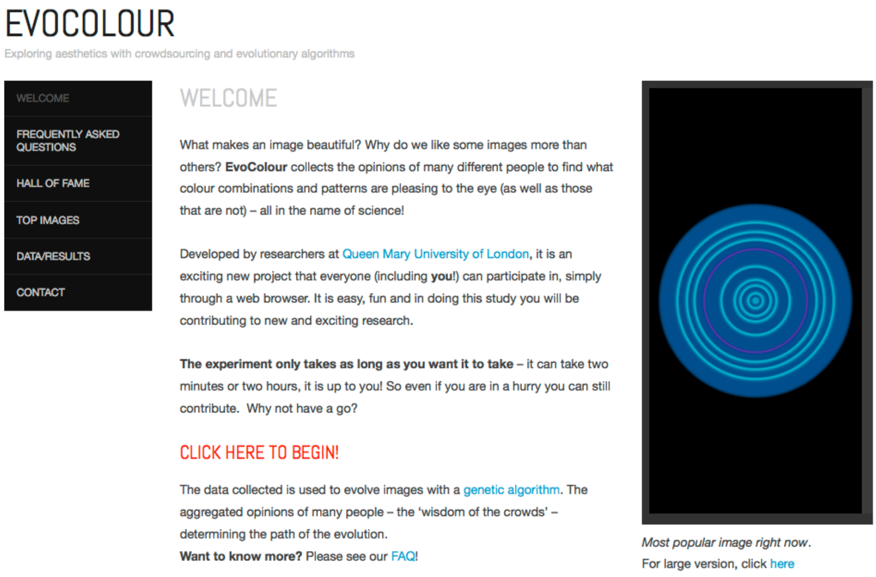


Figure 5.1: Screenshot of the *EvoColour* home page - www.evocolour.net

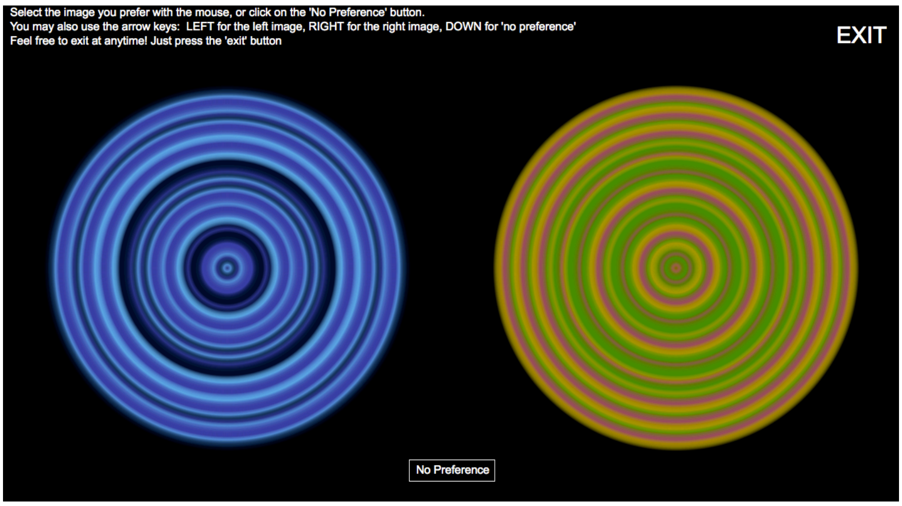


Figure 5.2: Screenshot of the *EvoColour* interface

which they could select a number of images to ‘keep’. The binary selection option was chosen here for two reason; the first is that it provides a low barrier to entry and engagement[2]. The second reason is that only rendering two images on screen ensures that images would occupy a large amount of screen real estate, allowing them to be rendered in high-fidelity.

The *genotype* of the images consist of a 3rd-order Markov process of three states, and three colours that map to each of the states. Additionally the values that are ‘fed’ into the Markov chain (the values that are traditionally provided by a random number generator) also formed part of the genotype. As such the *expression*, the process of rendering the *phenotype* (the image) from the genotype is deterministic. The genotype, expression, phenotype, *reproduction* (the image mating algorithm), *variation* (how the images mutate) and other characteristics of the evolutionary process are described in greater detail in section 5.4.

Matthew Lewis describes some of the considerations that go into using interactive EMA systems to elicit objective aesthetic preferences:

> “Researchers often write of a desire to collect information based on the user’s selection, and to mine this data for objective evidence of aesthetic preferences. Aside from very substantial problems of shifting selection context and attributing user intent, and the challenges of computational aesthetics practically speaking it has been very rare that sufficient usage data has been collected from these high-dimensional spaces to derive statistically significant aesthetic results”(M. Lewis, 2008).

One key distinguishing feature of *EvoColour* over other EMA systems is the restricted space of possible outputs. Having a small state space supports measuring and finding significant features, which in turn allows for a detail observation of the dynamics of the evolving populations. The images in *EvoColour*, and their evolution, lend themselves well to objective numerical analysis, making it possible to conjecture links to objective aesthetic preferences.

As will be made clear *EvoColour* did not find the parameters that yield the ‘archetypal image’, nor could it be stated with any certainty that evolved prominent feature combinations, and their presumed relationship to objective aesthetic preferences, could be generalised out with this specific context. However *EvoColour* does provide strong evidence that certain characteristics of colour values, combinations and arrangements are clearly preferred of others; helping uncover some aspects of the processes and mechanisms behind visual aesthetic preferences.

[2]In response to lessons learned from the *MelodyTriangle* mobile phone study.

Despite this restricted parameter space, there is nonetheless within it a considerable gamut of possible of images. In fig. 5.3 a number of randomly generated images are presented. Some images have regular repetitive patterns, others have random sequences, and are visually busy. Conversely some images are quite minimal - or even consist of just a single colour. These images are referred to henceforth as *Markov images*.

The features considered relate to colours and colour pallets, and the spatial arrangements of these colours. In the analyses, a large number of features are calculated on the images and the populations. Even with such a relatively small range of outputs, the gathered data is vast, and to extract the feature combinations that correlate with aesthetic preferences is not trivial.

In section 5.4 the experimental setup of *EvoColour* is outlined, including details of population sizes and mating processes. Three evolutions were run and with slightly differing mating algorithms, and with different levels of measurable success. Detailed analysis of the evolutions are provided in section 5.6, outlining the features that were deemed more popular (and those that were unpopular), and where appropriate, contextualising them in terms of previous research on colour preferences and design paradigms. This also includes an analysis of a non-evolving control population where the images were ranked by users in terms of popularity.

Section 5.7 contains a discussion of the overall findings, placing *EvoColour* in the broader context of research, and outlines what this study can tell us about design with generative processes.

There exists a large body of literature on human colour preferences and work on colour harmony. Additionally there is in the field of design a large swath of heuristics and conventions that form part of designers' practice and education. An overview of the previous research in colour preferences is provided in section 5.3. These these studies and findings are referred to in the analyses of the collected data. As will be shown, *EvoColour* will be able to support some of these, and not others.

The body of literature on colour theory spans centuries and is vast, from the works of Issac Newton via Goethe to the plethora of modern research in empirical aesthetics. As space limits considerations, only a very brief survey of developments colour theory and models will be summarised here.

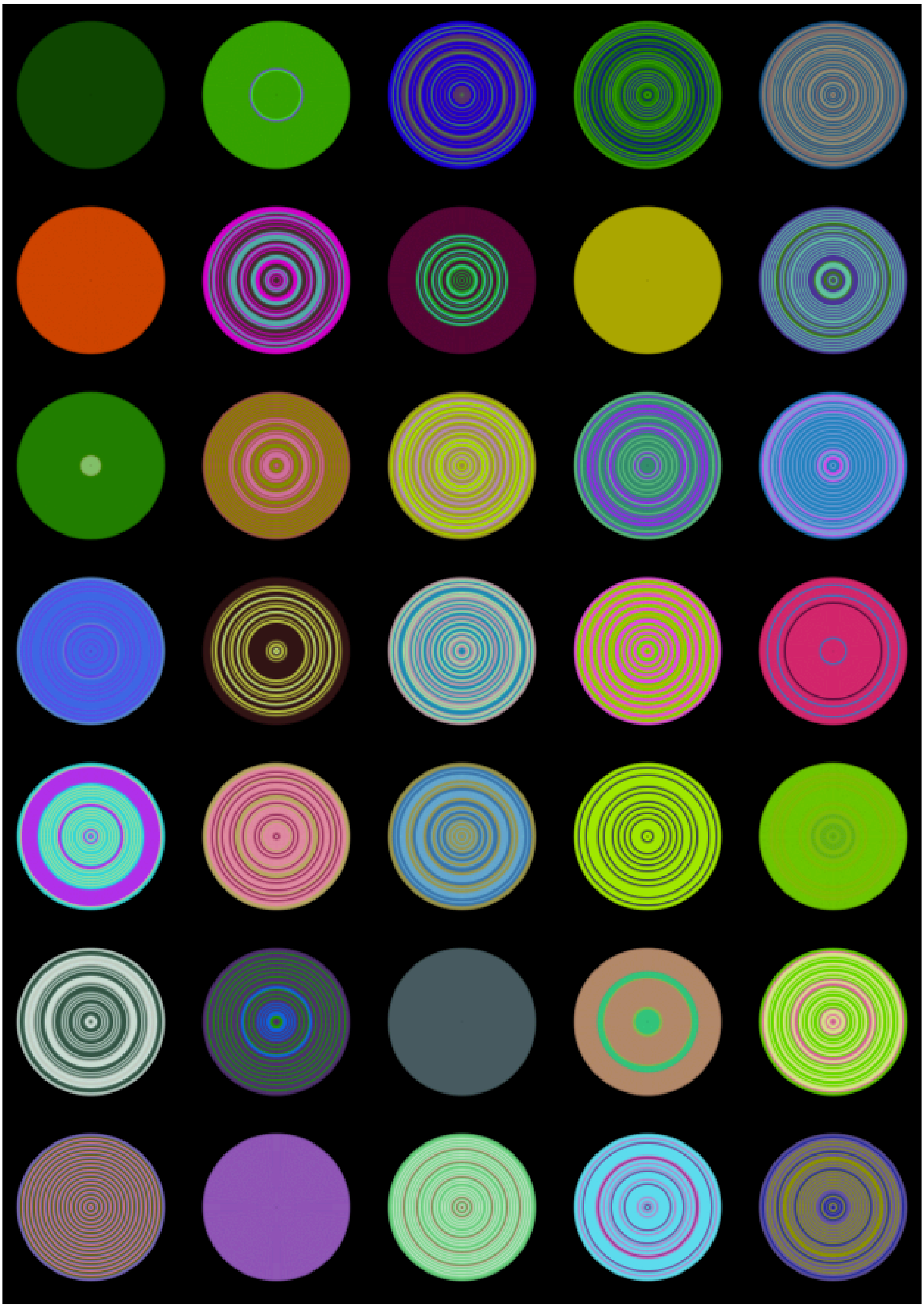
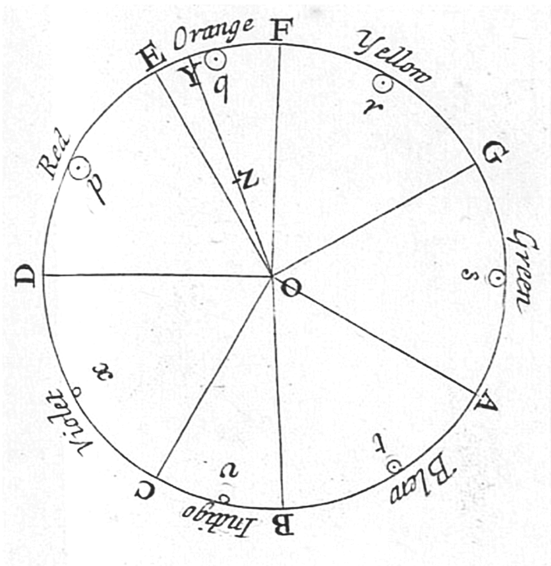
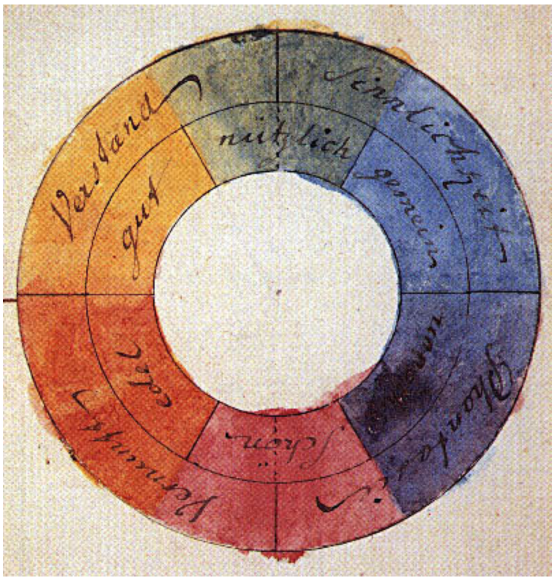


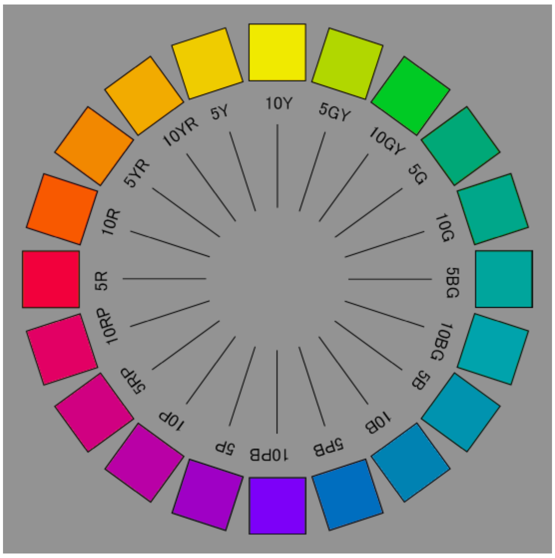
Figure 5.3: 35 randomly generated Markov images



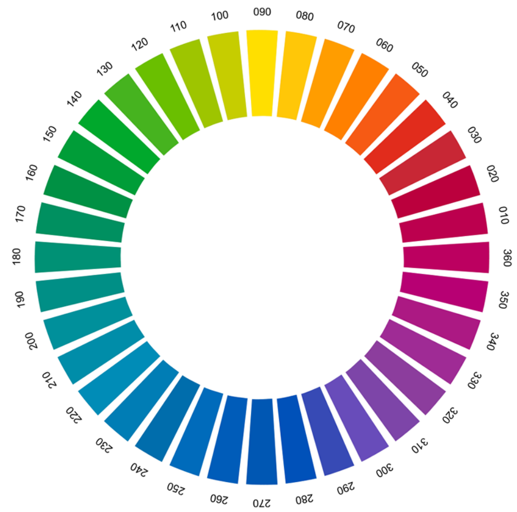
(a) Newton's Colour Circle



(b) Goethe's Colour Circle



(c) Munsell Colour Circle



(d) CIELAB hue colour circle

Figure 5.4: Colour Circles. The Munsell hue circle (c) and contains the archetypal hues that have often been used in empirical colour research (5B (blue), 5G (green), 5Y (yellow), 5R(red), 5P(purple)). In the CIELAB hue colour circle (d) hue values are in degrees, this is the system used in most modern colour research, and in the analyses of *EvoColour*.

(a) - By Isaak Newton [Public domain], via Wikimedia Commons
(b) - By Goethe, via Prof. Dr. Hans Irtel [Public domain], via Wikimedia Commons
(c) - By Jacob Rus (Own work) [CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0/)], via Wikimedia Commons
(d) - By Holger Everding (Own work) [CC BY-SA 4.0 (http://creativecommons.org/licenses/by-sa/4.0)], via Wikimedia Commons

5.2 Colour Theory and Colour Models

One of the first well know theories of colour was developed by Issac Newton, following from experiments 1660s, where he famously deconstructed white light to its component colours with a prism. From these findings he developed the first 'colour circle' introduced in his 'Opticks' in 1704, where he divided by seven the hues of visible light in a circle, mapping them to a musical scale(Newton, 1979). The colour wheel was then iterated upon and refined over centuries, and more intricate geometries of colour space were invented hand in hand with developments in paint mixing theory and technology.

In 1740 Castel defined a twelve-hue colour circle (again mapped to a musical scale, the chromatic scale in this instance), and in 1758 Tobias Mayer defined a colour triangle, with the primary colours on each side, subdivided in 11 intermediate colours. In 1772 J.H.Lambert opted instead to arranged colours in a 'colour pyramid'(Shevell, 2003).

Goethe responded to Newton's research in his 1810 'Theory of Colours', within he defined his own colour circle based on phenomenological considerations, rather than objective optical measurements, and complemented his circle with one of the first theories colour harmony. He defined colours as being harmonious if they were selected on from opposite ends of his circle(Von Goethe & Eastlake, 1970).

A significant breakthrough came in the early 20th century when Albert Munsell defined the *Munsell Colour System*(Munsell, 1912). Munsell identified three dimensions of colour that are still used in art education and colour research to this day: - *hue*, *lightness* (sometimes called *value*) and *chroma* (sometimes called *saturation*). This three dimensional colour geometry, with the cylindrical coordinates for hue and linear dimensions for lightness and chroma, as discussed in section 2.4.2, more accurately corresponded to a *conceptual space* for colour. Munsell colour system provides a standardisation of colours, and is often used to identify the colours used in empirical colour research, where the archetypal terms 'red', 'green', 'blue' 'yellow' and 'purple' often referring to 5 equidistant points in the hue circle, halfway up the lightness dimension, and at maximum chroma.

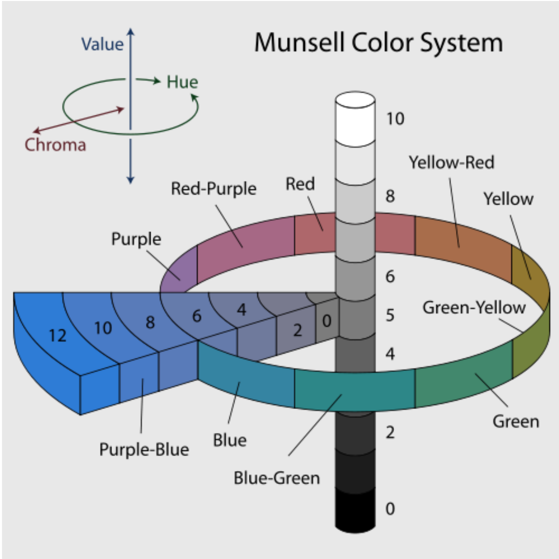


Figure 5.5: Dimensions of the *Munsell Colour System*. Lightness (or value) is on the vertical axis, chroma (or saturation) on the horizontal axis, and hue is the angle.
By Jacob Rus (Own work) [CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0/)], via Wikimedia Commons

5.2.1 CIELAB Colour Model

The *Munsell Colour System* is the basis (via a number of intermediate model) of the *CIELAB* colour model (or more formally *CIE* $L^*a^*b^*$[3]), used in both in the analysis, and the computations in the evolutionary algorithms of *EvoColour*. CIE stands for the *Commission Internationale de l'Eclairage*, a scientific organisation that has established standards for colour spaces since the 1930s.

$L^*a^*b^*$is a three dimensional colour space. The vertical L^*axis represents lightness. 0 has no lightness (i.e. absolute black) to 100 which is maximum lightness (i.e. absolute white). The a^*axis is green at the negative extreme, and red/magenta at the positive end. The b^*axis has blue at the negative extreme, and yellow at the positive.

CIELAB was designed to achieve perceptual uniformity, such that the euclidean distance, ΔE^*, of pairs coordinates in this space is representative of how different two colours are perceived to be. This colour-difference calculation is known as CIE76, however empirical studies have found this to not be as accurate as was intended. Subsequent standards, CIE94 and CIEDE2000 were then defined to correct irregularities in perceptual uniformity. In *EvoColour* the CIE94 standard(McDonald & Smith, 2008) is used to compute colour differences both in the analysis of results and in the evolutionary algorithm. CIEDE2000 was deemed to be too difficult to

[3]The asterisk are part of the full name, to distinguish from the *L a b* of yet another colour model, the Hunter colour model.

implement relative to the potential accuracy gain.

Any $L^*a^*b^*$coordinate can also be represented in cylindrical form, with hue angle and chroma axis through a simple transformation of the $a^*$and b^*coordinates to polar coordinates. This way of representing colour is known as $L^*C^*h°$. The C^*axis represents chroma or 'saturation'. This ranges from 0 at the centre of the circle, which is completely unsaturated (i.e. greyscale) to 100 or more at the edge of the circle for high chroma. The $h°$is the hue angle. The units are in degrees, ranging from 0°for red, 90°for yellow, 180°for green, and 270°for blue. This hue colour circle can be seen in fig. 5.4d.

$L^*a^*b^*$coordinates are 'device independent', and to represent an actual real life colour, a *white point* must be defined. There are numerous standards for white points, representing different lighting conditions. The CIE standard illuminant D65 was used as this is the standard white point for sRGB, the most common colour specification for computer monitors(Nielsen & Stokes, 1998). As participants in *EvoColour* view images on displays with differing capabilities and brightnesses, it is impossible to have a 'correct' white point defined. It is understood that the variety of brightnesses and capabilities of the participants devices would introduce noise into the data, but it is also expected that these difference would average out.

Relationship to RGB

When colours are to be rendered on a computer screen, they are most often represented as triplet values of red, green and blue (RGB values). sRGB is the most commonly used standard for rendering colour on electronic devices and screens.

The rendering of a pixel on screen, can in a sense be seen as a trivial generative process. And like so many other generative processes, there is an imperfect correlation of parametric similarity to phenomenal similarity. The Euclidan distance between two points in sRGB space has a poor correlation to perceptual similarity, unlike in $L^*a^*b^*$. In *EvoColour* the images have colour information stored as sRGB values, as that is how their colours are drawn to screen. However all colour transformations and analysis are done in $L^*a^*b^*$.

It is important to understand that $L^*a^*b^*$has a *gamut*, or range, of colours greater than that of human visual system; many coordinates are just 'imaginary' colours. Conversely sRGB has a significantly smaller gamut than all the colours perceivable by humans (i.e. there are many colours that can be seen, but that cannot be rendered on monitors and screens). Additionally sRGB can represent some areas of $L^*a^*b^*$space more densely than others. This is illustrated in fig. 5.6 and fig. 5.7, where thousands of randomly generated sRGB colours are plotted in $L^*a^*b^*$ and $L^*C^*h^\circ$ space.

To be able to perform accurate analyses of changes in colour distributions in *EvoColour*, it was necessary to identify the probability distributions of colours in $L^*a^*b^*$ and $L^*C^*h^\circ$ space when generated uniformly in sRGB. This is because the populations of *EvoColour* were seeded with randomly generated RGB values[4]. The distributions are presented in fig. 5.8. Note how in particular with regards to hue angle, h°, the distribution is far from even. It is thus vitally important, especially when considering the distributions of h° values image populations, to take this into account. These distributions are used as 'reference distributions' in the analyses of image populations to quantify where the shifts in colour occur. If colour had no effect on colour preferences whatsoever, it is expected that in the evolved populations distribution of colours would have a matching shape. By subtracting the distributions of the population under observation from these distributions, it makes it possible to identify the shifts, without being mislead by the natural tendency for the colour values to cluster.

[4]A future improvement would be to seed populations with colours evenly distributed in $L^*a^*b^*$ or $L^*C^*h^\circ$ space

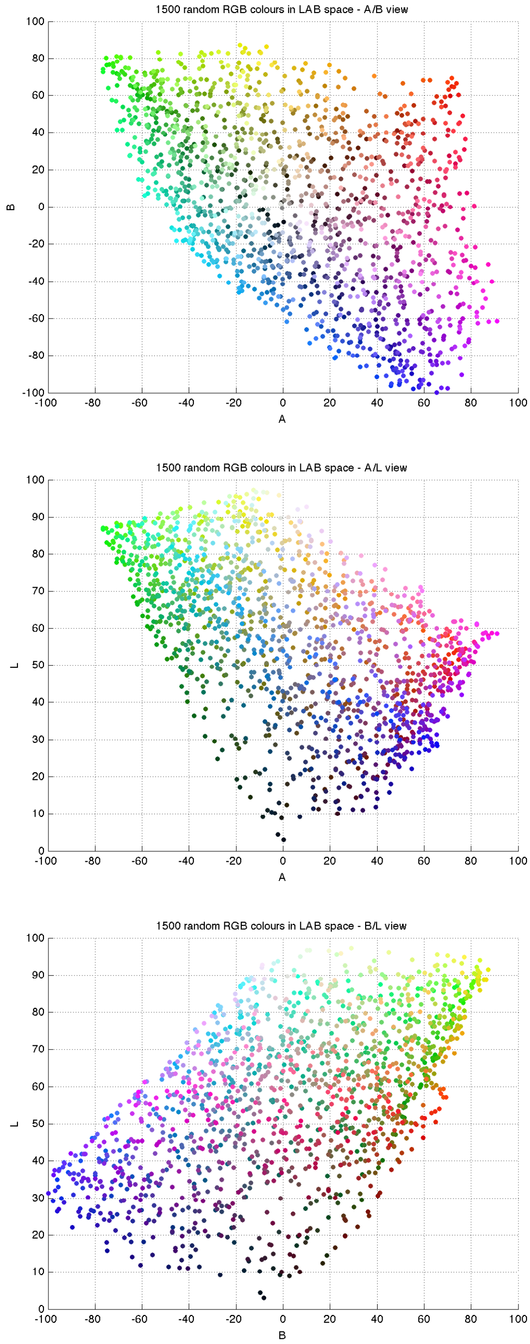


Figure 5.6: Distribution 1500 randomly generated sRGB colours in $L^*a^*b^*$space, a^*/b^*view, a^*/L^*view and b^*/L^*view respectively. Note how some areas of the space, such as areas with low L^*values are more sparsely populated.

Figure 5.7: Distributions of 1500 randomly generated sRGB colours in $L^*C^*h^\circ$space, C^*/L^*view, C^*/h°view and h°/L^*view respectively.

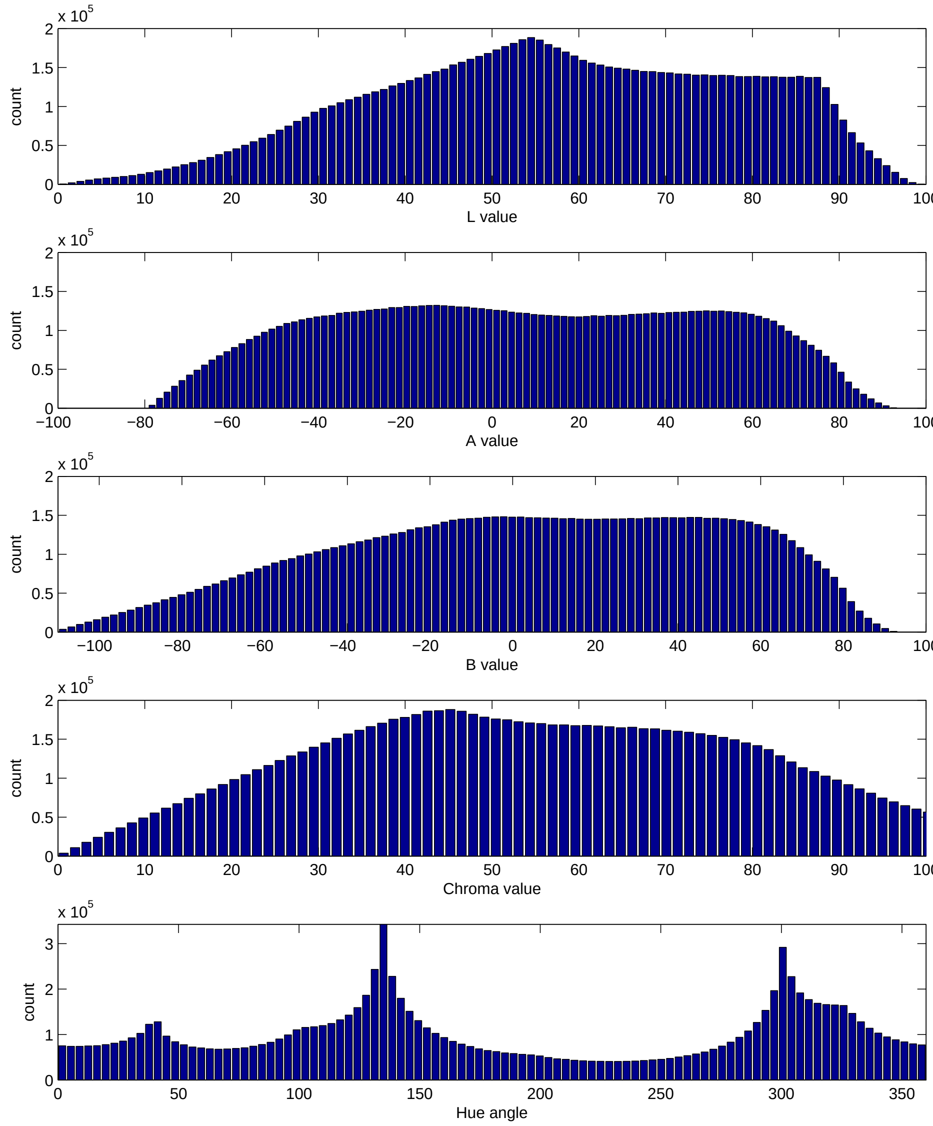


Figure 5.8: 100-bin histograms of 10 million randomly generated sRGB colours according to their L^*, a^*, b^*, $C^*$and h°values.

5.3 Previous Work on Colour Preferences and Harmony

5.3.1 Single Colour Preferences

Empirical research on colour preferences has been carried out since the late 1800s, and was predominently concerned with hue. Eysenk carried out an extensive survey of previous work on colour preferences in 1941, and established a ranking of hue preference in this order: blue, red, green, violet, orange and yellow(Eysenck, 1941). Other studies, such as some carried out by Granger(Granger, 1955) and Guilford and Smith(Guilford & Smith, 1959) in the 1950s considered colours of varying chroma and lightness as well, and found hue preferences in the order of blue, green, purple, red, yellow.

McManus et al.(McManus, Jones, & Cottrell, 1981) despite being very critical of the lack of rigour of previous work in empirical colour preferences studies, also find blue to be the most preferred hue, with yellow the most disliked. They state that there are however individual differences in preferences and the ranking of the intermediate hues are not as clearly defined as they vary with chroma and lightness, as well as with gender.

Crozier's more recent survey of hue preferences concluded that Eysenk's original ranking does on the whole hold true (Crozier, 1999). The most extensive single colour preference study was carried Palmer and Schloss in 2010, and reported that bluer colours are preferred over yellower colours, saturated (high chroma, mid lightness) colours are preferred over light colours and low chroma colours, and dark yellow and orange are the most disliked colours(Palmer & Schloss, 2010).

Other studies have identified different orderings relating to personality types (extroverts prefer brighter colours(Crozier, 1999)), gender (high proportion of males have blue as their favourite colour, whereas females are more evenly split between blue and green(Ellis & Ficek, 2001)), and cultural differences (Chinese observers liked more colours that were identified as 'clean' and 'fresh', where this tendency was not present for British observers(Ou, Luo, Woodcock, & Wright, 2004)), and colour associations (Stanford and Berkley students prefer the hues of their university's logos (Schloss, Poggesi, & Palmer, 2011)).

However the view that on the whole, blue is the 'world's favourite colour', is well established.

5.3.2 Colour Harmony

'Colour harmony', understanding how and when colours go well together, and are aesthetically pleasing, has also been a subject of research for centuries. Theorists have attempted to identify colour harmonies in terms of relative geometrical positions of the colours within a colour space.

As mentioned Goethe proposed an opponent colour harmony with colours harmonising if they are at opposite ends of the colour circle. Chevreul proposes two kinds of harmonies within his own colour system based on paint mixtures; 'Harmonies of Analogous Colours' and 'Harmonies of Contrasts'(Chevreul, 1855, p. 63). In the former, colours that are different lightness or chromas but of the same hue would harmonise, as would those that are of the same lightness and chroma, but are close to each other in the hue colour circle. In the latter, colours of contrasting lightness, contrasting hues, or in opposite ends of colour space in all dimensions would also harmonise.

Otswald suggested that colours would harmonise if they have equal black, white and hue contents in the colour space, a notion of balance(Ostwald, 1969). Itten suggested that hues would harmonise if they formed symmetrical polygons on the hue circle; 'dyads' for two colours,'triads' for three and so on, with similar contrastive geometries for other dimensions of colour space (Itten, 1961; Shevell, 2003). Munsell suggested that for colours to be harmonious, their amounts and positions should be balanced across one of a number of possible traversal paths cutting through the colour space(Shevell, 2003; Kuehni, 2002). Students of Munsell colour theory then could, given amounts of particular colours, programatically find additional colour/amount pairs that would lead to 'balance'. This is reminiscent of rules of harmony or counterpoint in music.

Art and design textbooks most often describe how to achieve colour harmony with reference to the hue circle(Westland, Laycock, Cheung, Henry, & Mahyar, 2007), with the most common geometries as follows:

- *Monochromatic harmony* - colours with the same hue
- *Complementary harmony* - colours from opposite ends of the colour circle
- *Analogous harmony* - colours close to each other in the colour circle

Additional geometries, such as Itten's 'triads' are also often presented. However the effect of lightness or chroma on the harmony of a hue pallet are less often considered(Westland et al., 2007).

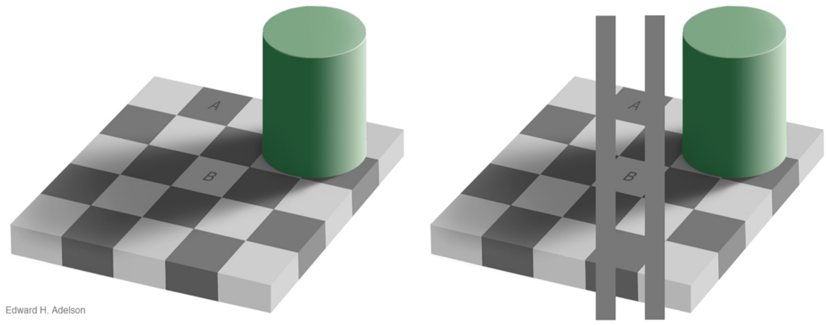


Figure 5.9: Adelson's Checker Shadow Illusion. The two squares *a* and *b* are of the same shade of grey.

The findings in empirical contexts are more vague and contradictory than the single colour studies(Schloss & Palmer, 2011). This is probably due to the numerous experimental variables that come into play when considering multiple colours. Each colour consists of three variables (hue,chroma,lightness), and the spatial arrangements and presentations of the colours (e.g. side by side, figure/ground) and their relative amounts all affect the outcomes.

Additionally there are a number of psychophysical effects and artefacts that muddy the issue; *contrast effects* can lead to the same colour to appear to be different based on the surrounding contexts (red on a blue background appears far more brilliant than when on an orange background). Further the human visual system is able to assign the same colour to an object, despite changes in illumination, a process known as *lightness* or *colour constancy*(Levine, 2000, p. 248). These effects are often demonstrated through visual illusions(Adelson, 1993), such as Adelson's checker shadow illusion (see fig. 5.9).

Moon and Spencer developed one of the earliest empirical attempts to arrive at a quantitative model of colour harmony. They extended Birkoff's previously discussed *Aesthetic Measure* to the colour dimension, with 'order' being a measure similarity and contrasts of hues, over the 'complexity' which is related to area, and finding the appropriate weightings experimentally(Moon & Spencer, 1944). Granger has carried out a number of studies in colour harmony in the 50s, and proposed that harmony increased the further apart the hues were in the colour circle(Granger, 1955), however these were not verified by more recent studies.

For instance Ou and Luo(Ou & Luo, 2006) have identified four principles that make colours pair harmonious, and contrasting hues is not amongst them:

In order of decreasing strength, they are:

1. *Equal hue principle* - Colours of the same, or very similar hue are judged to be harmonious

2. *High lightness principle* - Colour pairs that are of high lightness are judged to be harmonious.

3. *Moderate lightness principle* - Colour pairs that have either high difference in lightness, or small difference in lightness are harmonious (can be overruled by principles 1 and 2).

4. *Blue principle* - Colour pairs that are in the blue end of the chroma circle tend to harmonise (can be overruled by principles 1 and 2).

When considering the theories of harmony summarised above, Schloss and Palmer point out - "These theories are different enough that, if all their predictions were pooled, nearly every color pair could be considered harmonious!"(Schloss & Palmer, 2011). Further they provides possible explanations for the conflicting results in colour harmony empirical research. Notably, a lack of considerations for the distinction between 'harmony' and 'aesthetic preference' (disharmonious colours can still be 'preferred'), and a lack of consideration for the distinction between figure/ground comparisons (figural preference) vs side by side comparison (pair preference).

In their own studies, they found that hue similarity was the main factor in both pair preference and pair harmony, and found no increase in preference for contrastive hues. This supports the studies by Ou et al but conflicts with the findings of Granger, as well as with the predictions of many of the theorists, from Goethe, Otswald, Chevreul (the *Harmonies of Contrast*), and Itten ('dyadic' harmony). Further Schloss and Palmer found that 'preferred pairs' tended to have a large lightness contrast component, whereas 'harmonious pairs' have more similar hues and are lower in chroma. Interestingly, in distinguishing between side by side and figure/ground preferences however, they found some support for the 'harmonies of contrast' in the hue dimension : "warmer figures are preferred on cooler backgrounds, cooler figures are preferred on warmer backgrounds, and figures are generally preferred on backgrounds of contrasting lightness"(Schloss & Palmer, 2011). This indicates that geometric arrangements of colours to have a strong effect in how aesthetically pleasing the colour combinations are.

Few researchers have attempted to considered arrangements of more than two colours. A notable exception is the work of Hård and Svik, where they formulate a *Theory of Colors in Combination*(Hard & Sivik, 2001). It is a system for the descriptions of colour arrangements that can take into account the areas of size of individual colours, the perceptual similarities of the colours and the order 'rhythms' of the colours, however it does not carry with it any implications of the aesthetic value of the combinations, and is rather just a descriptive model.

In *EvoColour* the images consist of concentric circles of different thickness and varying sequences. Rather than specific sets colours, a *Markov image* can have all colours renderable by a computer monitor. Further the patterns arrangements of the colours, be they in a random sequences, or in predictable sequences (or somewhere in between), can all have an effect on overall aesthetic preference. A *Markov image* is a test material of a greater complexity than found in the experiments of empirical colour harmony research. (However as mentioned, it is simultaneously an object of much less complexity than found in other EMA systems.) It will be demonstrated that despite the greater complexity, significant trends in preferences can nonetheless be identified.

5.3.3 Hypotheses

Drawing on the survey of colour theory and empirical research above, a number of hypothesis that will be tested are presented in tables 5.1 and 5.2, and will be returned to in the results section 5.6.

	Hypothesis	Notes
1	*Strong preference for blue hues*	after (Eysenck, 1941), and others
2	*Strong dislike for yellow hues*	ibid.
3	*Preference for red hues*	ibid.
4	*Dislike for orange hues*	ibid.
5	*Preference for high chroma colours*	after (Palmer & Schloss, 2010)
6	*Dislike for light colours*	ibid.
7	*Dislike for low chroma colours*	ibid.

Table 5.1: Hypotheses relating to global colour preferences

	Hypothesis	Notes
8	*Preference for similar hues*	analogous/monochromatic harmony
9	*Preference for opponent hues*	complementary harmony
10	*Preference for equidistant hues*	triadic harmony
11	*Preference for large lightness contrast*	‘preferred pairs’ - after (Schloss & Palmer, 2011) -
12	*Preference for similar hues & lower chroma*	‘harmonious pairs’ - ibid.
13	*Preference for warm/cool figure/ground*	ibid.
14	*Preference for colours all of high lightness*	’high lightness principle’ - after (Ou & Luo, 2006)
15	*Preference for moderate lightness contrasts*	’moderate lightness principle’ - ibid.

Table 5.2: Hypotheses relating to colour combinations

Note that if some preferences are stronger than others, then the weaker preferences might be masked by the stronger ones in the evolutionary process. Further the empirical studies from which these hypotheses were derived considered equal areas of colour pairs. In *EvoColour* however the areas are not often the same, and there can be up to three colours present. As will be discussed, the analyses of the results in section 5.6 appear to support hypotheses 1, 2, 4, 6, 8, and 11.

5.4 Design

This section provides details of the design of *EvoColour*. It begins with the definition of the *Markov image*, and its genotype and expression into a phenotype in section 5.4.1. In section 5.4.2 the evolutionary algorithm behind *EvoColour* is described, including details of the reproduction (image mating) and image mutation processes, as are characteristics of the evolving populations, such as population sizes, mutation rates.

5.4.1 The Markov image

Genotype and Phenotype

The genotype of a *Markov image* consists of:

- A 3rd-order Markov chain of three states.
- Three colours that map to each of the Markov chain's three states.
- The *seeds*; the values that are 'fed' into the Markov chain.

The genotype of the image in fig. 5.10 is provided in table 5.3. The Markov chain is represented as a 3x27 matrix of probabilities, each determines the probability of going to a particular colour from the last three colours. The colours are represented as three values representing red, green and blue (the RGB value). As will be discussed later, all the matings and analyses are done in CIELAB colour space, not in RGB. The values that are 'fed' into the Markov chain, or *seeds*, are the numbers that would normally be provided by a random number generator in stochastic process. This is represented as a 50x3 matrix of values between 0 and 1, each row for each colour. Whenever a particular colour is selected by the Markov process, the next seed value is taken from the row of that colour.

In order to ensure that a mix of deterministic and non-deterministic Markov chains would be in the starting populations, a weighted random number generator, that increase the likely hood of 1s and 0s, was used in the creation of the Markov transition matrices.

The justification for this is that although a stochastic Markov process is used, it was deemed desirable to have a deterministic mapping from phenotype to genotype. As can be seen in

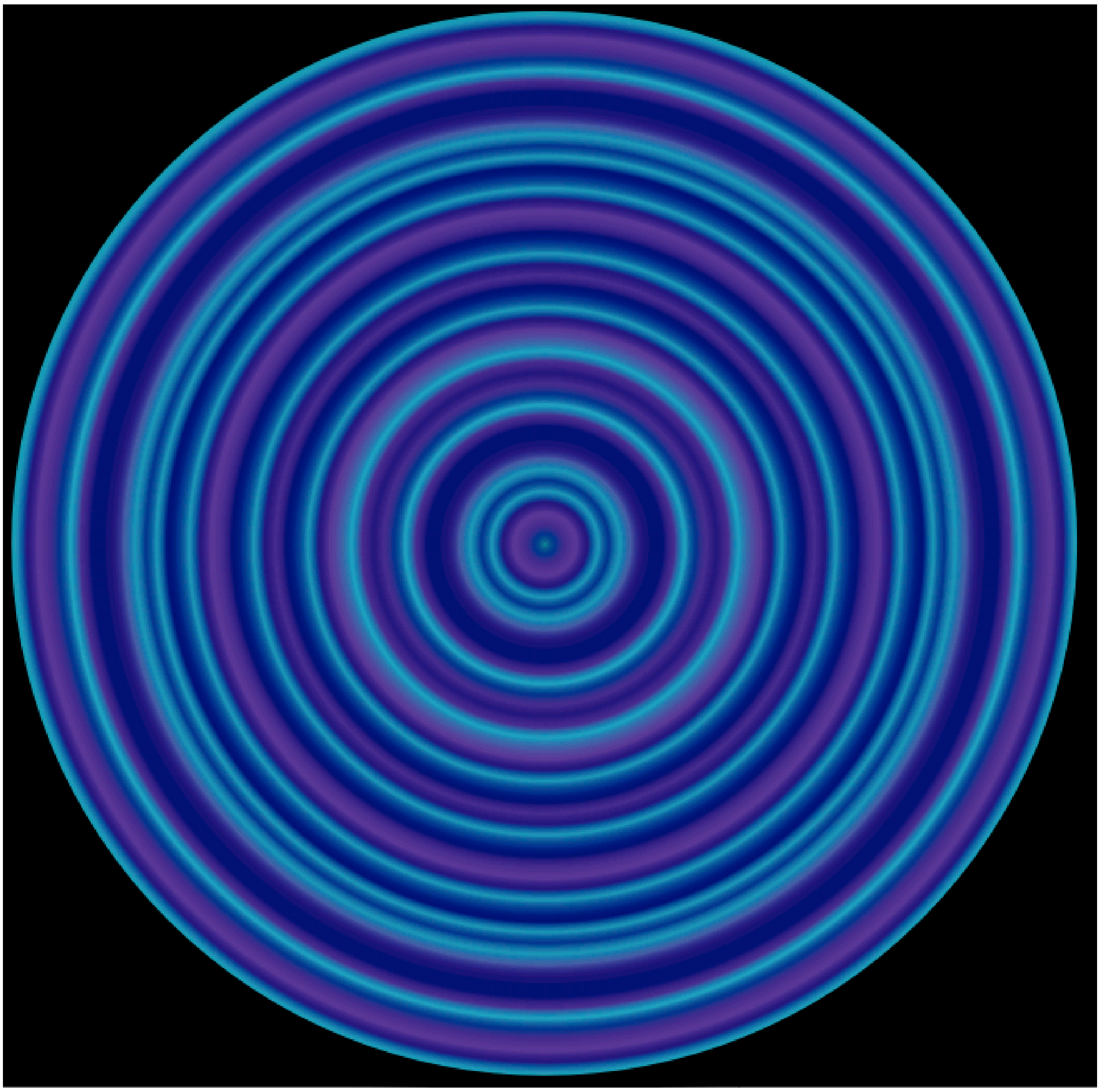


Figure 5.10: The *Markov image*

Table 5.3: The genotype of the *Markov image* in fig. 5.10. It consists of three elements, the *transition matrix*, the three colours, and the *seed values* fed into the markov chain. The transition matrix is represented as 3x27 matrix of probabilities. Each row represents the probabilities of going to a colour given the previous three colours, each row sums to 1. For example, given the previous three colours were *C A C*, the probability that the next colour is *A* is 0.27. Rather than using a random number generator to determine the next colour, a matrix of *seedvalues* is used, containing values between 0 and 1, each column belonging to a colour. Each time a colour is rendered, the next value input into the transition matrix is taken from the seed value matrix at the corresponding column. The output sequence (then mapped to the colours) for the image in fig. 5.10 is also shown. In bold are the portions of the sequence that are deterministic according to the transition matrix. For instance, once *A C C* has occurred, that *A B A* are the next symbols is deterministic.

Markov Transition Matrix

Next Colour ⇒	**A**	**B**	**C**
A A A	0.48	0.04	0.48
A A B	0.71	0.01	0.28
A A C	0.65	0.12	0.23
A B A	0.39	0.17	0.44
A B B	1.00	0.00	0.00
A B C	0.64	0.01	0.35
A C A	0.00	0.58	0.42
A C B	0.67	0.00	0.33
A C C	1.00	0.00	0.00
B A A	0.07	0.36	0.57
B A B	0.03	0.01	0.95
B A C	0.61	0.00	0.39
B B A	1.00	0.00	0.00
B B B	0.10	0.10	0.80
B B C	0.36	0.27	0.37
B C A	0.70	0.03	0.26
B C B	0.24	0.01	0.75
B C C	0.53	0.22	0.25
C A A	0.75	0.00	0.25
C A B	1.00	0.00	0.00
C A C	0.27	0.51	0.22
C B A	0.00	0.15	0.85
C B B	1.00	0.00	0.00
C B C	0.07	0.00	0.93
C C A	0.00	1.00	0.00
C C B	0.66	0.02	0.32
C C C	0.31	0.05	0.64
⇑ Previous Three Colours			

Colours (RGB)

	Red	**Green**	**Blue**
A (dark purple)	0.133	0.001	0.551
B (light blue)	0.214	0.961	0.934
C (light purple)	0.443	0.226	0.722

Seeds Values

50 x 3 matrix of floats (0 to 1)

Output Sequence:

B**ACCABA**BCAAACBACACBC**CABA**ACABA**ACCABA**ABABCAAACB**ACCAB**

fig. 5.11, images with the same colours and Markov chain can still look quite different from each other, yet they are clearly of the same 'family', and within a family some are clearly visually more compelling then others. Using Markov chains together with the seed values as phenotypes allows for specific images to be recreated. The Markov chain represents the overarching 'global' characteristics of the sequence, while the seed values carry information that can manifest itself in the details. Having a separate set of seed values for each colour allows for the seed values arrays to be spliced and re-combined in a genetic process, while still being able to pass on some visual characteristics.

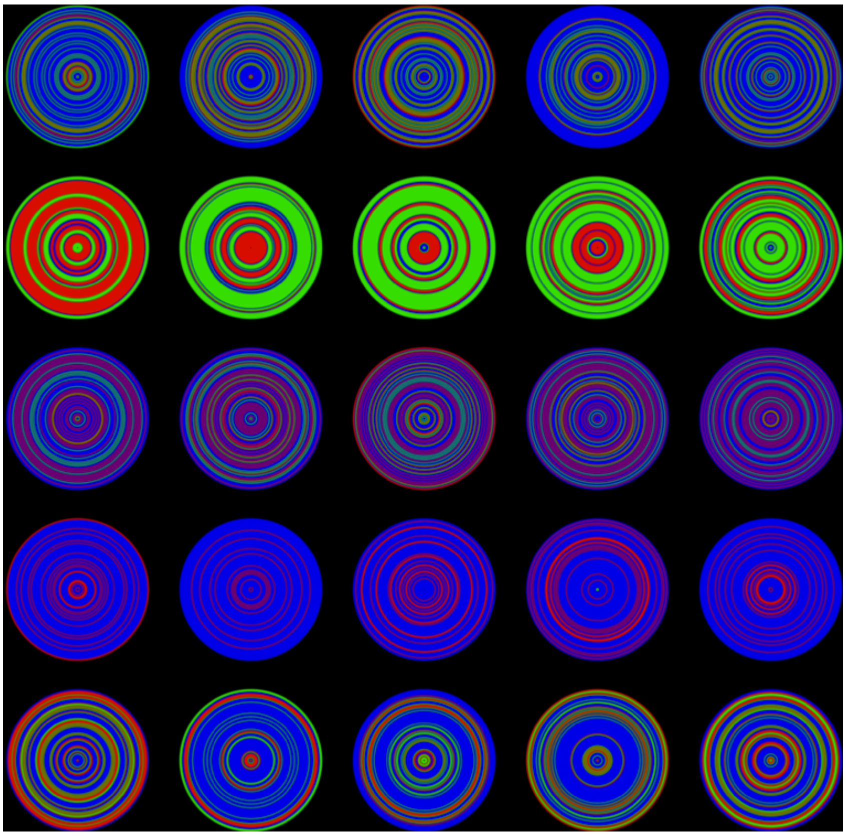


Figure 5.11: 5 'families' of Markov images. Each row has the same Markov chain and colours, but do not share seed values.

The process of creating the image from the genotype consisted of two steps. The first steps generates the output sequence out of the Markov chain and seed values. The second step is taking this sequence of symbols and associated colours values, and beginning in the middle drawing circles of increasing radii until all 50 were rendered.

The implementation of these two steps were separated for technical reasons (see 5.4.4), the first on the server side and the latter, the drawing, was done on the client side (the user's web browser). However there is no conceptual separation between these steps, they are both just part of the same process of expression of the genotype into the phenotype.

There is slight cross-fade between adjacent colours. Admittedly this could potentially skew results, however it was deemed that the slight fade made for considerably more pleasing image renderings. As maintaining user interest and engagement is of primary concern to the success of interactive evolutionary processes, it was deemed to be a risk worth taking.

State Spaces

Before outlining the details of the evolutionary process and algorithms, it is worth considering further the spaces in which the *Markov images* reside.

Beginning with the space of all possible images renderable on a screen, the relationship to a Markov image can be characterised like this:

- The space of all possible images drawable on a screen *is greater than*
- The space of all possible images of three colours *is greater than*
- The space of all possible images of three colours renderable with concentric circles *is equal to*
- The space of all possible three colour images renderable with concentric circles of colours as generated by a Markov process.

Note how in this case, the space of the last two categories are of the same maximum dimensions, because any sequence of symbols could be output by a Markov process. However the key point is that the *probability* of finding any one particular image is very different depending on the process used to generate the output sequence. It is far more likely to find an image with a repeatable sequence of circles if the implementation is driven by a Markov process, then if it there is no such process, and there was an equal likelihood of any one of the three symbols/colours after any other[5].

The generative Markov process then can be understood as having the effect of changing the probabilities of a particular sequences, making it more likely that a particular kind of pattern appears over another. Anyone particular Markov process, bar the ones that have unit probabilities (of 1 and 0), can generate any sequence of symbols.

As a demonstration, in fig. 5.12 is a sequence of images that all have the same Markov chain of high entropy. With a very simple computational fitness function – where the more blue present

[5]Equivalent to a Markov process that has equal probabilities of going from anyone state to the next, the transition matrix is filled with values 0.33. As will be discussed later, due to a bias in the mating algorithm towards images of maximum entropy, the first runs of *EvoColour* were of images of this type.

in an image, the 'fitter' it is – the seed values input into the chain are randomly mutated. The input sequences that improve the fitness score are kept, and this is repeated until the image evolves to one of a maximum fitness. This shows how any one Markov chain, and by analogy any one generative process, does not necessarily reduce the space of possible outputs - it is still possible for this Markov chain to generate any image within the space of concentric circles - however this particular chain is far more likely to create an output image of high entropy that it is one of zero entropy.

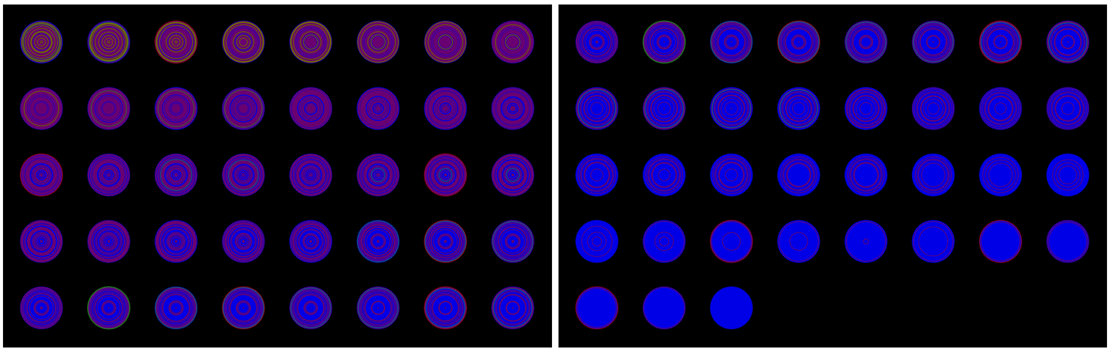


Figure 5.12: Markov images of the same, high entropy Markov chain with seed values mutated towards maximum blueness.

Similarly, a higher-order Markov process is capable of generating sequences that are possible, but unlikely in a first-order Markov process. A sequence such as *A-C-B-B-A-C-B-B-A-C-B-B-A-C-B-B-A-C-B-B-A-C-B-B* is not one that can be deterministically generated by a three state first-order Markov process. This is because there needs to be a state, a memory, lasting longer than one symbol; it must be possible to say 'after C,B, got to B with 1.0 probability', and 'after B,B go to A with 1.0 probability'. Given an output of finite length however, a run of a first-order Markov process can generate this sequence, however not reliably.

This further illustrates that any one generative process embodies not only a space of possible outputs, but also embodies probabilities and likely-hoods of particular outputs. It would have been possible to design *EvoColour* with images without any Markov chains, mating images simply by splicing the sequences of symbols directly. However a higher level coherence and structure would be much harder to come by as the likelihood of images with random sequences would be much higher. Operating at this higher level of Markov processes allows for a much more rapid exploration of the possible phenomenal range of the outputs.

5.4.2 Evolutionary Algorithm

Process

In developing an evolutionary algorithm, numerous decisions with regards to implementation and parameter values need to be made. This includes finding parameters for population sizes, chances of mutation, number of selections before a new generation, wether parent images remain in the population after mating or not, as well as the various alternatives methods (and *their* parameters) of the mating processes. It is difficult to know how to find ideal values for these parameters.[6] This is a *meta-parametrisation*, a process that in most EMA works are selected in an ad-hoc fashion by the artist, with a combination of experimentation and intuition. In the literature on EMA works, even though the choices for these parameters are often documented, it is not as common to read the processes by which they were selected or the motivation for the selections described.

For *EvoColour*, these were found by way of semi-automated process. A number of computational *fitness functions* were developed, each representing a feature that could possibly be related to aesthetic preference. The fitness functions included preferences for images of certain global average colours along all the axes of LAB space, distances between colours in colour space, as well as the 'null-fitness function', where the virtual user would select images at random. Moreover a 'multi-objective fitness function' was developed, that could embody any number of the above fitness functions at any ratio of preference. For instance one multi-objective function could be - *70% prefer images that are bright, and 30% prefer images that have colours far apart in colour space*. A complete list of the computational fitness functions developed can be seen in table 5.4.

A batch process then ran numerous evolutions, each with one virtual 'user' that would embody one or more of these fitness functions. Selections would happen instantly, and so many evolutions of hundreds of generations could be run in a few minutes. With this semi-automated framework, different parameters for the evolutionary algorithm were tried out and tested, as well as different implementations of the image mating algorithm. It was monitored how quickly the populations would evolve towards images that fulfilled the fitness functions. By ensuring that the populations could evolve towards any and all of the fitness functions, the parameters for the evolutionary algorithm were thus selected, as was the implementation of the mating algorithm.

[6]The irony, of course, is that often the purpose of the evolutionary process itself is to find parameters.

Fitness Function	Reinforces
averageColour_Max_LAB_L()	High L^*average image colour
averageColour_Min_LAB_L()	Low L^*average image colour
averageColour_Max_LAB_A()	High a^*value (reds)
averageColour_Min_LAB_A()	Low a^*value (greens)
averageColour_Max_LAB_B()	High b^*values (yellows)
averageColour_Min_LAB_B()	Low b^*values (blues)
averageColour_Max_LCH_C()	High C^*values (saturated colours)
averageColour_Min_LCH_C()	Low C^*values (greyscale)
averageColour_LCH_H(H)	Images with h°at H
maximumLABDistanceBetweenColours()	Colours far apart in $L^*a^*b^*$space
equalLABDistanceBetweenColours()	Colours equidistant in $L^*a^*b^*$space
maximumEntropy()	Maximum entropy in colour sequence
matchLAB(target)	Images close to *target*
null()	Random images
multiObjective(fitnessFunctions,weightings)	Combinations of fitness functions

Table 5.4: Computational fitness functions used to test the evolution parameters and implementations

The *null fitness function*, were the virtual users would select images at random, was also run to ensure that the mating algorithms didn't carry within them too strong a tendency to pull the population towards certain areas of the solution space. The type and amounts of mutation was determined by monitoring that the populations would not have a tendency to homogenise too quickly and resist genetic drift, while still being able to evolve the images towards their fitness goals.

As it turns out all mating algorithm implementations and evolutionary algorithm parameters attempted had a tendency of pulling the population towards certain areas of features over others (genetic drift), and also have a tendency towards homogenisation of the images. Genetic drift and homogenisation was counteracted with mutation. In this way, it was ensured that the final algorithms' drifting tendency was much weaker than the genetic force of selections.

This meta-parametrisation process is very similar to the meta-parmetrisation process in the *Keyebernates* system, described in chapter 4, or the 'ecosystemic' generative art and music pieces described in section 2.2.4. It is a process a-kin to 'tuning', finding a way for the system as a whole, including the end user(s), to reside in the range between noise and stasis; to not end up being completely static, nor ending up behaving completely randomly. This in-between space allows volition to guide the overall system in an overall direction, without the system being stuck in a certain area of solution space[7].

[7]It is perfectly reasonable to consider finding the parameters of the evolutionary algorithm themselves with an automated evolutionary process. The process above could be completely automated, however then this meta-genetic algorithm would itself need to have its own parameters that would themselves be

When *EvoColour* was launched, users would alternatively make selections on three populations of images, each time being shown two randomly selected images from the population. This included a non-evolving *control population* - referred to henceforth as *CP* - and two evolving populations - referred to henceforth as $EP1$ and $EP2$. The difference between $EP1$ and $EP2$ was that parent images would be removed after mating in $EP1$, but would remain in the population for $EP2$. Additionally each user would have their own 'local evolutions' that would evolve independently only with selections from that user, however it was quickly noticed that on average there would not be enough selections for these local evolutions to go beyond a couple of generations, and so this feature was subsequently disabled.

In addition to selections on the populations, every fourth selection, users would do a 'population comparison', where they would compare an image from the first and the latest generation of the populations. This was done to be able to track the overall fitness improvements of the evolving populations.

Despite the apparent relative rigour in choosing the implementations and parameters values for the evolving populations, one omission led to the first release of *EvoColour* to have a tendency, regardless of user selection choices, to evolve towards images of maximum entropy. This was because in the process described above, there was no computational fitness function defined that would prefer images of *minimum* entropy. As such when the Markov chain mating process had an inherent bias towards generating chains of increased entropy, this was not detected until after the analysis of the first round of public evolutions was carried out.

This was then corrected, and a re-launch of *EvoColour* was made with a new evolving population - $EP3$. However the first evolutions still provide an interesting case study. $EP1$ and $EP2$ still evolved colour preferences, but for a different generative process; that of images who default to randomly arranging the three colours.

The final parameters values, selection and generation counts for *CP*, $EP1$, $EP2$, and $EP3$ are displayed in table 5.5.

In the next section the image mating algorithms are described.

5.4.3 Image Mating

A core aspect of evolutionary systems is the process of combining the genotypes of parents to create new genotypes of offsprings - the mating algorithm. How well the relevant phenomenal

discovered, and so on, ad-nauseam.

	Number of Images	Selections/ Generation	Mutation Probability	Elitist	Evolves to Maximum Entropy	Number of Generations	Number of Selections
Control Population (*CP*)	1000	n/a	n/a	n/a	n/a	n/a	15775
Evolving Population 1 (*EP*1)	500	100	5%	No	Yes	70	6974
Evolving Population 2 (*EP*2)	500	100	5%	Yes	Yes	70	6973
Evolving Population 3 (*EP*3)	500	100	5%	Yes	No	82	8104
						Total:	37826

Table 5.5: Attributes of populations in *EvoColour*

attributes of the parents are transmitted to the children is one of the key determinants of the success of an evolutionary algorithm. This is sometimes referred to as *transmission fidelity*, and has been cited as a limiting factor to the performance of evolutionary algorithms (e.g. (MacCallum et al., 2012)) as well as the longevity of socially transmitted cultural knowledge(H. M. Lewis & Laland, 2012). Intuitively, one would expect the children of two parents to have phenomenal attributes of both parents, and averaged over a population, the phenotype strength - the average fitness - should go up when the fitter members of the population are mated.

As such, the design of a mating algorithm should be able to support this transmission of phenomenal attributes from parents to offspring, and also do so in a way to minimise the chance that overall fitness decreases. One way of doing this is to incorporate knowledge of what kinds of phenomenal attributes are the ones that ‘matter’ for overall fitness in the way genotypes are combined in a mating algorithm.

For automatic evolutionary processes, where fitness is not determined by human users, but by a computational process, the rate of selections possible is far greater. This means that the mating algorithms can afford to be much looser and the ratio of children who actually improve the fitness can be far smaller. However in interactive evolutions, where the fitness is determined by infinitely slower human evaluators, it is of primary concern that the designer of the mating algorithm has some awareness of what might consist features of value.

Mating Colour Palettes

As described in section 5.4.1, the colour pallets of a *Markov image* consists of three colours. Mating the colour pallets of two parents can be defined as:

Given colours of Parent 1 ($P1_{c1}$,$P1_{c2}$,$P1_{c3}$), and the colours of Parent 2 ($P2_{c1}$, $P2_{c2}$, $P2_{c3}$), for any number of Children, generate three colours (Cx_{c1}, Cx_{c2}, Cx_{c3}).

The first approach considered is analogous to the crossover technique of standard genetic algorithms. Each colours of the parents was viewed as a ‘gene’ to be passed to the children. For each of the three colours, a child would inherit one of the colours of the parent. For example, one child might end up with ($P1_{c2}$, $P2_{c1}$,$P1_{c3}$), as its colours, and another might end up with ($P2_{c1}$, $P1_{c1}$,$P2_{c2}$)

Although clearly some phenomenal attributes of the parents are passed on to the children, this approach did not perform well in the batch evolutions run with the computational fitness functions described in the previous section. There are multiple reasons for this, the first is that this approach does not take into consideration the *amounts* of colour present in each parent. In *Markov images*, not all three colours are always visible; sometimes the phenotype has two or even just one colour. If all colours in the genotype are treated as equals, a single-colour parent could pass on to its children colours that are not visible at all, and in an extreme case, it is possible that a child has only colours that are not visible in either parent.

Another consideration is that this mating algorithm only considers the colours’ absolute values, and does not consider the relationships *between* the colours. As discussed in the overview of colour research in section 5.3.2, the geometric relationships between colours in colour space is of relevance when considering wether colour pairs are harmonious or preferred[8]

When this ‘naive’ algorithm described above was tested against the computational fitness functions, the populations were unable to optimise quickly towards the fitnesses that considered the relative positions of colours in colour space as valuable.

As such, a more sophisticated colour mating algorithm was designed for *EvoColour*, that embodied two key considerations: global average colour (as reflected in studies of single colour preferences in section 5.3.1), and the geometric relationships between the colours in colour space (as reflected in the studies of colour harmony, described in section 5.3.2). Although there were

[8]Although what the geometric relationships that yield harmonious colours *are* is debated. However being able to transmit the relationships from parent to child can help ensure that any colour harmony is potentially passed on.

multiple conflicting findings in the study of colour harmony, they usually described colour harmony as a geometry in colour space.

The first step of the algorithm converts the phenotypes' values from RGB into $L^*a^*b^*$values prior to mating, so as to maintain as close a link to the phenomenal attributes of the images as possible. The final step of the mating converts the $L^*a^*b^*$colours back to RGB for storage and rendering.

A 'centre of mass' in the colour space of an image is calculated. This 'centre of mass' takes into consideration the amounts of each colour that is visible in the image, calculated by weighting the colours in the genotype based on the number of circles of that colour rendered. Then three 3-dimensional vectors in $L^*a^*b^*$representing the distance and direction in colour space between each of the colours is calculated. These vectors represent the relationships of the colours have to each other, and can be understood as a defining triangular plane in the $L^*a^*b^*$colour space.

This algorithm attempts to transmit the geometrical relationships between the colours from the parents to the children. When mating, child images inherit the centre of mass (the average colour) of one parent, and the triangular shape (the colour relationships) of the other. This is akin to translating, or shifting, the triangle in colour space to another area of $L^*a^*b^*$colour space. An example of this is seen in fig. 5.13, where two parent colour pallets swap the centre of mass and their colour relationships.

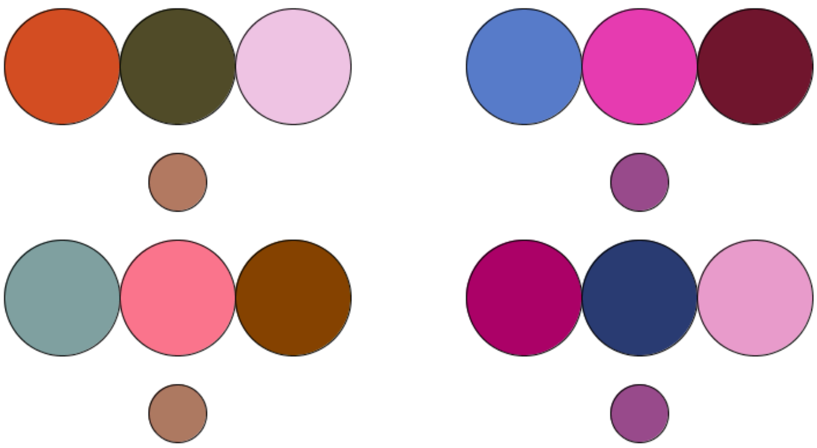


Figure 5.13: An example of the colour palette mating in *EvoColour*. The top row represents the parents, and the bottom row are two children. The 'centre of mass', the average colour in $L^*a^*b^*$is represented as the small circles. The centre of mass is passed down to the child directly below each parent, but the colour geometry is passed to the other child.

In this process it is possible that one of the vertices of the triangle falls outside the gamut of colours renderable in RGB. If this occurs, the algorithm attempts to rotate the triangle so that all three points fall in the gamut of RGB, while keeping the same centre of mass, maintaining the relative distances between the points of the triangle. If it is impossible to rotate the triangle, then the colour is clipped to the edge of the RGB colour space, and how much clipping occurred is stored as an 'error' value. If this 'error' falls above a certain threshold, then the algorithm falls back on the naive colour swapping algorithm described earlier. The amount of error tolerated was found with the semi-automated process described in 5.4.2.

This error can be seen as a form of *mutation*, a certain amount of which is actually desirable in genetic algorithms. Because of the possibility that vertices of the triangle might fall outside the gamut of renderable colours, numerous repeated iterations of this algorithm had a tendency of squeezing the triangle into smaller shape. However this was offset by mutation processes described below.

This colour mating algorithm performed much better at fulfilling the criteria of the set of computational fitness functions, listed in 5.4, then the naive algorithm, which was unable to significantly improve the population in terms of a preference for certain colour relationships and distances. This is because the design of the algorithm has embedded within its design the knowledge that colour relationships, and not just absolute colour values, are of relevance to the final aesthetic value of the output.

Mating the Markov Chain

The first design of the mating algorithm for the Markov chain part of the genotype was a 'blend crossover' technique, a common approach for continuous-value parameter interpoltation. Here the floating point parameters, in this case the values in the Markov transition matrices, are 'cross-faded'. However this technique carries within it the assumption that parametric proximity is tied to phenomenal similarity. This turned out to not be the case when considering a Markov transition matrices.

This technique works fine if one chain is of high entropy (say all the values in the chain are 0.3), and the second Markov chain is deterministic (say all the values in the chain are either 1s or 0s). Any 'crossfade' of these matrices would yield a Markov chain that would have predictability between those of the parent matrices. However this technique becomes a problem if the two parent matrices are of low entropy; crossfades of such chains would go *via* high entropy. This issue is why evolving populations *EP1* and *EP2* pulled images towards high entropy (see fig. 5.14).

When the Markov chains have reached maximum entropy, then it is equivalent to rolling a dice before picking the next colour. However structure could still be passed along in genotypes via the seed values that were fed into the Markov chains.

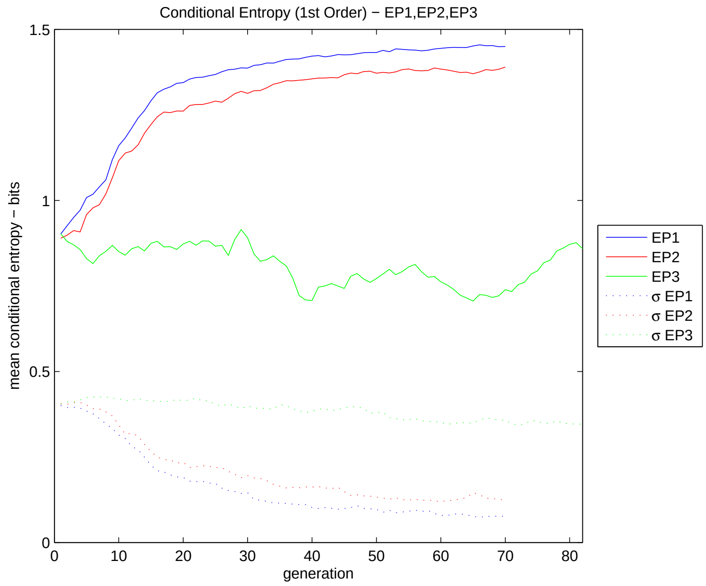


Figure 5.14: The first-order conditional entropy of $EP1$, $EP2$ and $EP3$. As can be seen, $EP1$ and $EP2$ tend towards maximum entropy quickly, whereas the 'corrected' evolution $EP3$ does not have this increasing trend. Note how $EP2$'s rate of entropy increase is somewhat lower than that of $EP1$. This is due to the elitism helping slow the bias towards high entropy.

In the 'corrected' *EP3*, children simply inherit the Markov transition matrix of one parent or the other.

Mating the 'Seed' Values

The 'seed' values – the numbers that are fed into the Markov chains when generating the output sequence – are mated with a traditional genetic crossover technique. A randomly selected 'crossover point' is selected that would cut the sequence of both parents at the same point. The children would then have a combination of the parents' sequences from either side of the crossover point.

Mutation

Mutation is valuable in evolutionary algorithms to ensure that new information enters populations, and also to resist premature homogenisation and to counteract genetic drift. The probability of mutation was determined by process described earlier. As such there was an 5% chance for anyone image to undergo a mutation in the mating process. If a mutation was set to occur, then the image could either have a mutation in colour, where one of the colours in the genotype would be changed to a randomly selected new colour, or in the seed values, whereby a whole new sequence of values would be generated. These mutations would make the resulting image have many of the same attributes of the original image yet still be noticeably different.

Elitism

One of the design decisions of evolutionary algorithms is wether to make the evolution 'elitist'. In an elitist evolution, the selected parents are allowed to stay in the population after mating. In *EvoColour* elitism was tested in the first two evolutions $EP1$ and $EP2$, where the former was not elitist, and the latter was. In preliminary analyses, it was determined that $EP2$ was able to better resist the bias of the flawed Markov chain mating (see fig. 5.14). When it came to the relaunch it was therefore decided to make $EP3$ elitist, as this was evidence that elitism would help counteract any further biases in the mating processes.

5.4.4 Implementation Details

EvoColour was developed with a client/server model. The server side of the implementation was developed with Java based technologies, supported by a MySQL database. The server side would store all the data, run the mating algorithms and generate the output sequences. The phenotype would be transferred to the client side - the users web browser - encoded as three colours and a sequence of 50 symbols for rendering. The rendering would be run in a Processing.js app, a javascript port of the popular Processing platform, embedded in the web-page.

EvoColour is hosted on the Amazon cloud, running on an ElasticBeanstalk Java container for high reliability and scalability. Data analysis consisted of exporting the contents of the Markov image into JSON objects, and then imported into Matlab for the quantitative analysis.

5.4.5 Dynamics of User Interaction

When users begin the study, they are presented with a short registration form, where they were asked to choose a username. Additionally they are asked for their age and gender, as well as being asked two survey questions relating to their interest and experience in the visual arts[9]. To provide as few barriers to entry and engagement as possible, all questions asked were optional. Once a user has registered, a cookie would be stored on the user's web browser containing just

[9] It was intended to use these survey questions in the analysis of individual 'mini-evolutions' local to each user. However the mean user selection counts were too low, so they were discontinued.

their username, allowing them to return to the site and continue to do more selections without needing to re-register or remember their username.

As of 23rd May 2015, there were 615 individual users, with a total selection count of 45499. A summary of user information is provided in table 5.6.

User count	615	**Gender**	
Total selections	45499	**Male**	249
Mean Selections/User	74	**Female**	288
Median Selections/User	32	**No Answer**	77
Max selections	1179		
Age Mean/StdDev	29.6 / 10.6		

Table 5.6: *EvoColour* user statistics as of 23rd May 2015.

A gamification element was included in *EvoColour* to help increase engagement and the number of selections, in the form of a 'Hall of Fame' of users with the most selections. A screenshot of the 'Hall of Fame' is provided in fig. 5.15. When a user exits they are told their ranking, and it is suggested to them that they could do more selections to climb up the rankings, as well as spread awareness of *EvoColour* via social media widgets. In hindsight, it is quite likely that significantly more user engagement would have been achieved if the 'gamification' element was made more explicit. This could have been done by indicating the current user ranking on screen during selections, and the number of selections needed to achieve the next ranking.


HALL OF FAME

Most Prolific Users

USERNAME	SELECTION COUNT
colourfix	1179
amber	1166
qwertie	826
jossie	762
yyyyyyyyyy	759
norm	709
aj9	637
gerarg	637
annefrench	626

Figure 5.15: The *EvoColour* 'Hall of Fame' as of 23rd May 2015.

Overall statistics were carried out on the selections, and it was found that there were more selections of the right hand image over the left, as seen in fig. 5.16. As images would be randomly assigned to the left or right, this will have had no impact on the result. Right-hand selections were slightly faster than the left hand selections, as seen in fig. 5.17.

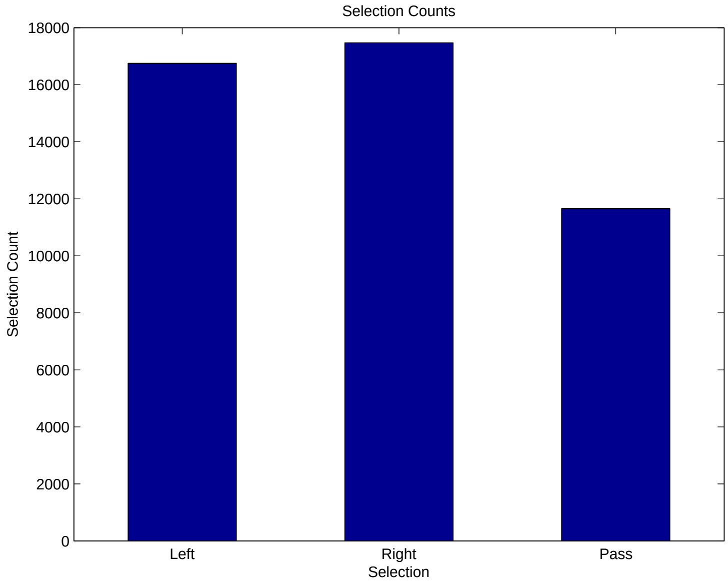


Figure 5.16: Selection counts for left image, right image and 'no preference' respectively.

5.5 Analysis Approach

Features Extraction

In order to carry out analyses on the populations of *EvoColour*, numerous features were calculated on the images in a batch automated process. Theoretical considerations of features and their interpretations were discussed in section 2.4.1 of the background chapter. A complete list and description of features is provided in table 5.7. It is important to note that the features were calculated on the output images directly, as oppose to being derived from parameters in the genotype.

The features include general properties such as the number of visible colours, information theoretic measures such as the conditional entropy and mutual information of the output sequences, and overall colour features, such as the mean $L^*a^*b^*$values. The colour features were also calculated for different sections of the images, dividing the sequence of colours in three. The first third representing the innermost colours, and the last third representing the outermost colours. In addition to the features relating to absolute colour values, features were calculated on

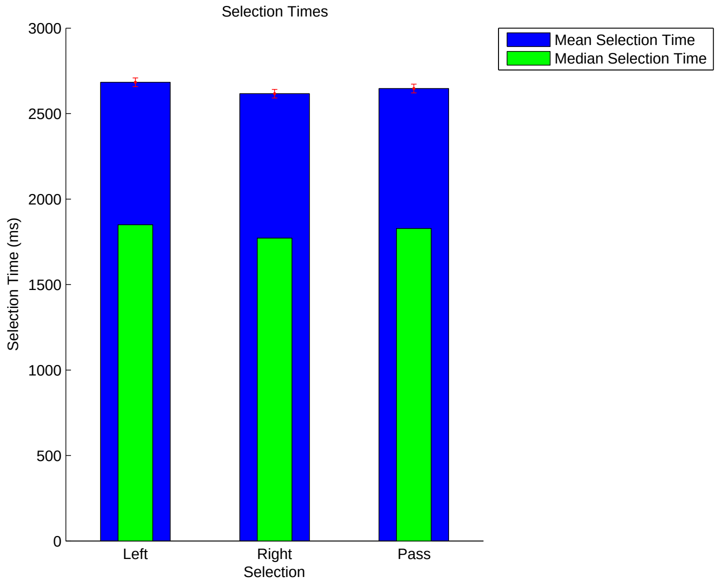


Figure 5.17: Mean and median selection times of left image, right image and 'no preference' respectively

the *change* in colour within an image. These include features such as $\mu\Delta C^*$, the mean change in C^*between consecutive colours, and $\max\Delta(\Delta E^*)$, the maximum $L^*a^*b^*$distance between consecutive colours. This is a total of 95 features on each *Markov image*[10].

In addition to these feature calculations on individual images, the mean and standard deviations of all the features in table 5.7 were also made on the populations of images as a whole, amounting to 190 features per population. This allows overall properties of evolutions to be tracked on a generation per generation basis.

Feature Selection

Given so many features, it is naturally challenging to identify find the features of relevance when characterising how the populations evolve and when searching for trends in aesthetic preferences. In this analysis both manual and computational approaches[11] to feature selection were carried

[10] Many additional features were also calculated, including features in RGB colour space, as well as features based on higher moments of distributions, such as *skewness* and *kurtosis*. However these turned out to be very noisy and so were omitted from the analysis.

[11] *Principal component analysis* (PCA) were also extensively carried out on the populations of images to help elucidate the relationships between the variable features. However the *PCA* analyses were not so successful. It is not clear why that was the case, but it is postulated that two key reasons may lie behind this. The first is that the co-variance between the variable features are non-linear and complex, and the other is that the data, in addition to being quite noisy also possessed a significant number of outliers, which is difficult to manage in PCA analysis. Some overall trends could be identified in with PCA. These indicated that the greatest variance between generations, and within populations, related principally to changes in the amount of blueness (b^*). However not many other insights were gained over other approaches taken. Therefore the details of the PCA analyses will not be included here.

out. Manual tests of specific features values where carried to answer specific questions, such as testing some of the hypotheses outlined in section 5.3.3.

In the discussion on self-organisation in section 2.5 of the background, it was described how the data-collection of the experiments could be understood as a decrease in relative entropy; the system increases in internal order by taking in structured information from the environment ('the users'). The feature selection reflects this by seeking to quantifying to what degree each feature accounts for the increase in 'self-organisation'.

As the analysis progresses a number of *observations* will be noted, and will be returned to in the final discussion in section 5.7. The first analysis considered is that of the control population, *CP* in the following section.

Table 5.7: List of *Markov Image* features.

Feature Symbol	Description
General	
$\|C\|$	The number of colours in the image
$\max\|\vec{c}\|$	The longest unbroken run of single colour in the image
$\|\Delta c\|$	Number of colour changes
Information Features	
H_X	Entropy (unordered)
$H_{Y\|X}$	Conditional entropy
$H^2_{Y\|X}$	Conditional entropy, 2nd-order (colours grouped in twos)
$H^3_{Y\|X}$	Conditional entropy, 3rd-order (colours grouped in threes)
$I_{Y;X}$	Mutual information
$I^2_{Y;X}$	Mutual information, 2nd-order
$I^3_{Y;X}$	Mutual information, 3rd-order
Overall Colour Features	
$\mu(L^*)$	Mean L^*value
$\mu(a^*)$	Mean a^*value
$\mu(b^*)$	Mean b^*value
$\mu(C^*)$	Mean C^*value
$\mu(h^\circ)$	Mean h°value
$\max(L^*)$	Max L^*value
$\max(a^*)$	Max a^*value
$\max(b^*)$	Max b^*value
$\max(C^*)$	Max C^*value
$\min(L^*)$	Max L^*value
$\min(a^*)$	Max a^*value
$\min(b^*)$	Max b^*value
$\min(C^*)$	Max C^*values
$R(L^*)$	Range of L^*values
$R(a^*)$	Range of a^*values
$R(b^*)$	Range of b^*values
$R(C^*)$	Range of C^*values
$\sigma(L^*)$	Standard deviation of L^*values
$\sigma(a^*)$	Standard deviation of a^*values
$\sigma(b^*)$	Standard deviation of b^*values
$\sigma(C^*)$	Standard deviation of C^*values
Localised Colour Features	divides image in 3
max, min, mean, range of colour features above.	$1/3$ = innermost $2/3$ = halfway $3/3$ = outermost
e.g.: $\max(C^*)_{1/3}$	e.g. Maximum chroma in innermost third of image
e.g: $\mu(b^*)_{3/3}$	e.g. Mean b^*value in outermost third of image
Change in Colour Features	
$\max\Delta(\Delta E^*)$	Maximum $L^*a^*b^*$distance (ΔE^*) between consecutive colours
$\max\Delta L^*$	Maximum change in L^*value between consecutive colours

Feature Symbol	Description
$\max\Delta a^*$	Maximum change in a^*value between consecutive colours
$\max\Delta b^*$	Maximum change in b^*value between consecutive colours
$\max\Delta C^*$	Maximum change in C^*value between consecutive colours
$\max\Delta h^\circ$	Maximum change in h°value between consecutive colours
$\min\Delta(\Delta E^*)$	Minimum $L^*a^*b^*$distance (ΔE^*) between differing consecutive colours
$\min\Delta L^*$	Minimum L^*change between differing consecutive colours
$\min\Delta a^*$	Minimum a^*change between differing consecutive colours
$\min\Delta b^*$	Minimum b^*change between differing consecutive colours
$\min\Delta C^*$	Minimum C^*change between differing consecutive colours
$\min\Delta h^\circ$	Minimum h°change between differing consecutive colours
$\mu\Delta(\Delta E^*)$	Mean $L^*a^*b^*$distance (ΔE^*) between differing consecutive colours
$\mu\Delta L^*$	Mean change in L^*value between differing consecutive colours
$\mu\Delta a^*$	Mean change in a^*value between differing consecutive colours
$\mu\Delta b^*$	Mean change in b^*value between differing consecutive colours
$\mu\Delta C^*$	Mean change in C^*value between differing consecutive colours
$\mu\Delta h^\circ$	Mean change in h°value between differing consecutive colours
$R\Delta(\Delta E^*)$	Range of $L^*a^*b^*$distances (ΔE^*) between consecutive colours
$R\Delta L^*$	Range of change in L^*value between consecutive colours
$R\Delta a^*$	Range of change in a^*value between consecutive colours
$R\Delta b^*$	Range of change in b^*value between consecutive colours
$R\Delta C^*$	Range of change in C^*value between consecutive colours
$R\Delta h^\circ$	Range of change in h°value between consecutive colours

5.6 Results

This section outlines the quantitative results of *EvoColour*. Section 5.6.1 provides an analysis of the control population *CP*. Section 5.6.2 describes the measures of fitness improvements in *EP*1, *EP*2 and *EP*3. Sections 5.6.3 and 5.6.4 contain analyses of the of evolving populations, together with evaluations of the hypotheses derived from the literature in colour preference and harmony.

5.6.1 Analysis of Control Population

The control population, *CP*, consisted of 1000 images, and had a total of 15775 selections made, each from randomly selected pairs of images. This population did not evolve, instead images would accumulate a score which was the number of selection 'wins' minus selection 'losses' (a 'no preference' selection would not alter the score). The images were ranked based on their score.

The distribution of image scores can be seen in fig. 5.18, and characteristics of the rankings can be seen below in table 5.8. All the images of *CP* are provided as item *Ex*4 of the illustrative materials accompanying this thesis[12], and are ranked according to popularity.

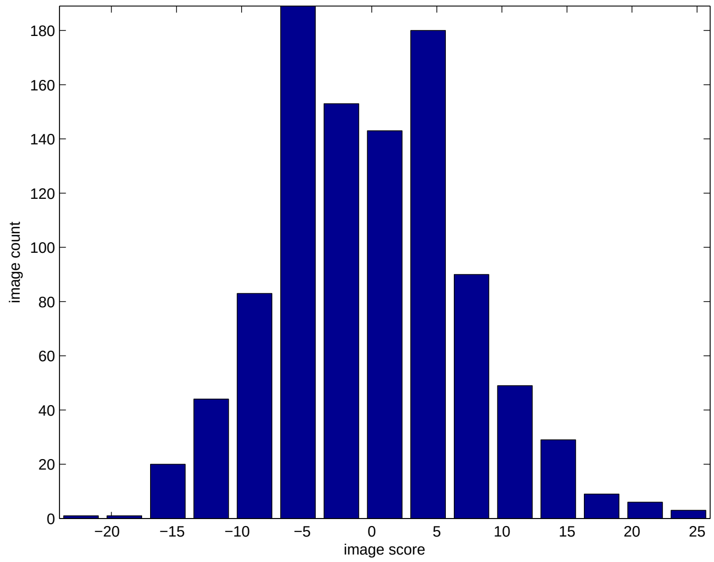


Figure 5.18: Image score distribution for *CP*

max score	26
min score	-24
mean score	0
standard deviation	7.30
skewness	0.23

Table 5.8: Image score distribution attributes for *CP*

The top 20 images of *CP* can be seen in fig. 5.19. The bottom 20 can be seen in fig. 5.20. Considering the top 20, it is interesting to note that these images look significantly different from each other. There is a variety in overall average colour, some have matching hues, others seem to have contrastive colours, and there are many differences in the geometric arrangements and patterns of the circles. Considering the bottom 20, the most obvious thing to note is that it is predominately populated by single colour images.

[12]The images can also be seen online:
https://www.flickr.com/photos/140600562@N05/collections/72157663025467944/

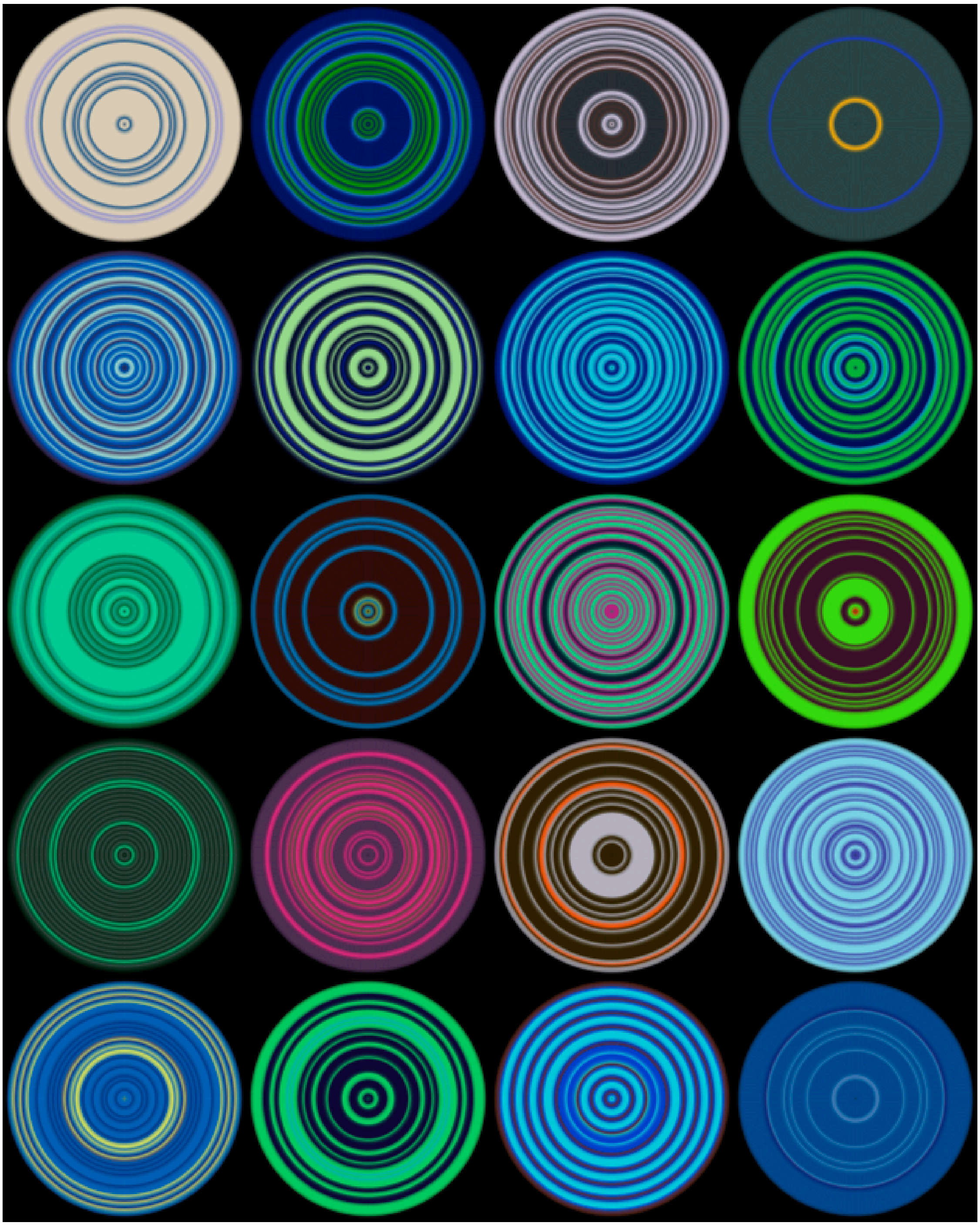


Figure 5.19: Top 20 images in *CP*. Top left is the image with highest score.

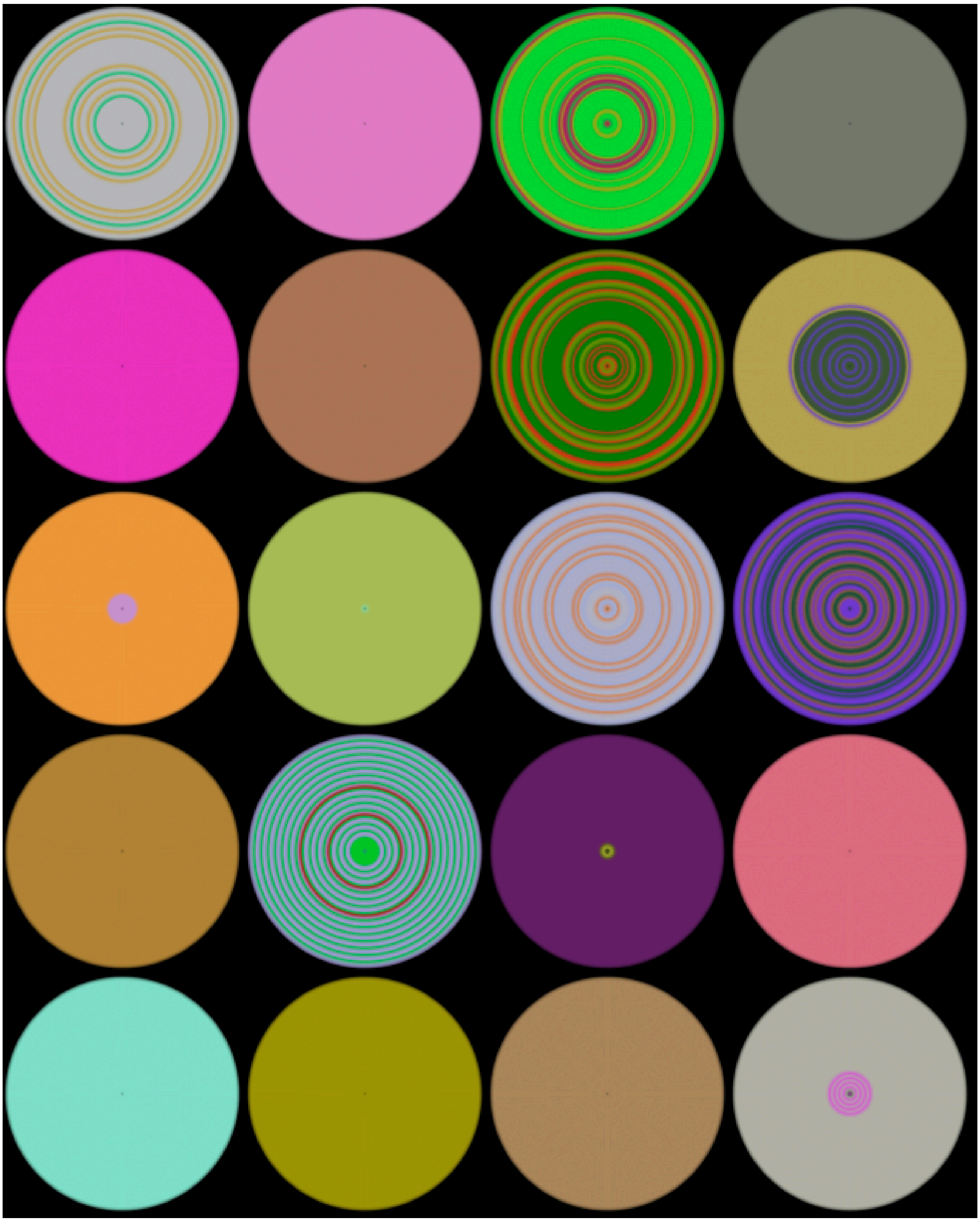


Figure 5.20: Bottom 20 images in *CP*. Bottom right is the image with the lowest score.

To identify which features are the ones most responsible for the ordering is not easily done by eye, perhaps with the exception of the easily identifiable unpopularity of single colour images. As such a computational method had to be employed to identify the relevant features.

This was approached by using the image score as a 'weighting' in calculating distributions of values for each feature. If an image had a score of say, +10, then its feature values would be counted ten times in building a distribution of 'positives'. If an image had a score of -5, then its feature values would be counted five times in building a distribution of 'negatives'. (Images of score 0 would not be included.)

It is then possible to see how a feature's values change with image popularity, by taking the *difference* of the 'positives' distributions and the 'negatives' distributions. The result can be displayed in a histogram with both positive and negative values, a number of examples of which will be displayed later in this section. Such plots will be referred to as *correlation histograms*.

An example is shown in fig. 5.21. The y-axis, *shift amount*, reflects how much this feature value is likely to shift with image score. This plot indicates that having only a single colour in the image has a strong negative correlation with image preference, and that having three colours has on average a weak positive correlation with preference. Having two colours, has a weak negative correlation. The maximum magnitude of the bins histogram is 1.0, this would indicate that features at this value are only present in positively ranked images.

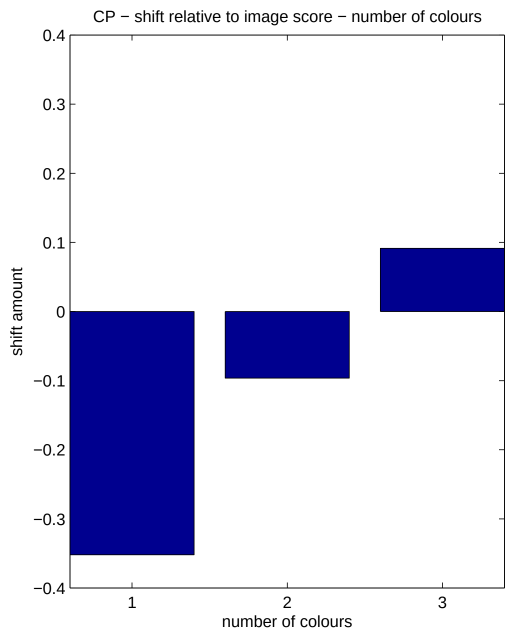


Figure 5.21: Correlation histogram for the number of colours, $|C|$, in *CP*

A value – *total shift* – estimating how different the *shape* of the ‘positives’ and the ‘negatives’ distribution are from each other, was calculated for all features. The *total shift* was calculated by quantising the ‘positives’ and ‘negatives’ cumulated feature values into 10-bin histograms, normalising the histograms, taking their difference, and then calculating the absolute magnitude of this difference.

Sorting the features by this ‘amount of shift’ returns a ranking, describing which features’ value distributions change the most with image score. As such they quantify how much the feature was subject to the organisation imposed by the image ranking. Note that this value does not indicate in what way the feature values change with ranking, just the amount they change relative to each other.

In preliminary analyses, it was observed that correlation histograms often have a large negative peak at values that corresponded to the values held by single colour images. This is because, as it has already been established, single colour images were very unpopular. It would thus prove more insightful to disregard single colour images when making these feature relevance estimates.

This ranking of features is shown in table 5.9. A *Abs Peak* value is provided, which represents the magnitude of the largest (positive or negative) change within the 10-bin correlation histograms. The further down the list of features one progresses, the noisier the correlation histograms plots become, and the more cautious one must be when attributing significance to the feature orderings.

Table 5.9: Features of images in *CP* sorted by *total shift*, which represents how much the distribution of that features' values change with respect to image ranking. *Abs Peak* is the absolute magnitude of the largest feature value shift in a 10-bin *correlation histogram*. Single colour images are omitted.

Feature	Total Shift	Abs Peak
$\min(L^*)_{1/3}$	2.14	0.35
$\max \Delta L^*$	2.06	0.40
$R(L^*)$	2.04	0.35
$R(L^*)_{1/3}$	1.91	0.35
$\mu \Delta L^*$	1.88	0.36
$\min(L^*)_{3/3}$	1.88	0.30
$R(L^*)_{3/3}$	1.88	0.35
$R(L^*)_{2/3}$	1.86	0.37
$\min(L^*)_{2/3}$	1.77	0.30
$\min(L^*)$	1.76	0.25
$\mu(h^\circ)_{3/3}$	1.67	0.42
$\mu(h^\circ)$	1.65	0.38
$\mu(h^\circ)_{2/3}$	1.55	0.30
$\mu(h^\circ)_{1/3}$	1.55	0.39
$\sigma(L^*)$	1.49	0.35
$\max(b^*)$	1.17	0.24
$\mu(b^*)_{3/3}$	1.15	0.30
$\mu \Delta(\Delta E^*)$	1.13	0.23
$\mu(C^*)_{2/3}$	1.09	0.23
$\min(b^*)_{3/3}$	1.09	0.21
$\min(C^*)_{2/3}$	1.08	0.24
$\max(L^*)_{2/3}$	1.06	0.21
$\mu(C^*)_{3/3}$	1.04	0.23
$\min(C^*)_{3/3}$	1.01	0.24
$\mu(b^*)_{1/3}$	0.97	0.19
$\max \Delta(\Delta E^*)$	0.96	0.16
$H^2_{Y\|X}$	0.93	0.30
$\max(L^*)_{3/3}$	0.92	0.21
$\min(a^*)_{2/3}$	0.91	0.14
$\mu(L^*)_{3/3}$	0.90	0.27
$\max(L^*)_{1/3}$	0.88	0.23
$\min(C^*)_{1/3}$	0.87	0.21
$\min(C^*)$	0.86	0.19
$\max(L^*)$	0.86	0.20
$\min(a^*)_{3/3}$	0.85	0.22
$\mu(C^*)_{1/3}$	0.83	0.17
$\mu(C^*)$	0.82	0.20
$\mu(b^*)$	0.82	0.19
$I^3_{Y;X}$	0.77	0.16
$\mu(L^*)$	0.77	0.18
$\min(b^*)_{2/3}$	0.76	0.18
$\max(b^*)_{1/3}$	0.75	0.22
$R(a^*)_{2/3}$	0.74	0.20
$R(a^*)_{3/3}$	0.73	0.18
$R(C^*)_{2/3}$	0.72	0.22
$\max(h^\circ)$	0.72	0.21
$R(b^*)_{2/3}$	0.71	0.25
$\mu \Delta h^\circ$	0.71	0.13
$\mu(L^*)_{1/3}$	0.70	0.16
$R(C^*)_{3/3}$	0.69	0.21
$\max(b^*)_{2/3}$	0.69	0.15
$\min(a^*)_{1/3}$	0.69	0.11
$H_{Y\|X}$	0.67	0.18
$\|\Delta c\|$	0.66	0.19
$\mu(a^*)_{1/3}$	0.66	0.15
$\min(b^*)$	0.65	0.13
$\mu(a^*)_{3/3}$	0.65	0.23
$I_{Y;X}$	0.64	0.11
$\mu(L^*)_{2/3}$	0.64	0.13
$R(b^*)_{3/3}$	0.63	0.18
$H^3_{Y\|X}$	0.62	0.29
$\max(b^*)_{3/3}$	0.62	0.17
$I^2_{Y;X}$	0.62	0.13
$\sigma(C^*)$	0.61	0.15
$\max(C^*)_{2/3}$	0.60	0.12
$\min \Delta(\Delta E^*)$	0.56	0.18
$\min(a^*)$	0.54	0.11
$\min \Delta L^*$	0.54	0.23
$\mu \Delta b^*$	0.52	0.12
$\max(C^*)$	0.52	0.11
$\mu \Delta C^*$	0.51	0.12
$\sigma(a^*)$	0.51	0.12
$\mu(b^*)_{2/3}$	0.51	0.17
$\max \Delta h^\circ$	0.51	0.09
$\mu(C^*)_{3/3}$	0.49	0.10
$\min(b^*)_{1/3}$	0.49	0.15
$R(a^*)$	0.49	0.08
$\sigma(b^*)$	0.48	0.12
$\mu(a^*)_{2/3}$	0.47	0.14
$\max \Delta C^*$	0.46	0.11
$\max(C^*)_{1/3}$	0.46	0.11
$\max \Delta a^*$	0.43	0.09
$\max(a^*)$	0.43	0.11
$R(C^*)$	0.40	0.09
$R(a^*)_{1/3}$	0.40	0.11
$R(b^*)_{1/3}$	0.39	0.09
$\max(a^*)_{2/3}$	0.38	0.09
$\max \Delta b^*$	0.38	0.10
$\max(a^*)_{1/3}$	0.37	0.11
$R(C^*)_{1/3}$	0.37	0.11
$\min \Delta C^*$	0.36	0.11
$\max(a^*)_{3/3}$	0.35	0.09
$\max\|\vec{c}\|$	0.34	0.15
$\min \Delta b^*$	0.34	0.12
$R(b^*)$	0.34	0.07
$\mu(a^*)$	0.34	0.10
$\mu \Delta a^*$	0.33	0.06
H_X	0.33	0.10
$\min \Delta h^\circ$	0.26	0.06
$\min \Delta a^*$	0.22	0.05
$\|C\|$	0.14	0.11

As can be seen, the 9 highest ranked features relate to L^*, image lightness. However they are not features describing the mean lightness of the image, this feature $\mu(L^*)$is much lower in the rankings, at position 40. They either represented the minimum lightness value, or relate to the change of lightness *within* an image.

The top ranked feature $\min(L^*)_{1/3}$, is the *minimum* lightness in the middle of the image. The other minimum lightness features are also highly ranked. The correlation histogram for the minimum lightness $\min(L^*)$is provided in fig. 5.22. Single colour images are omitted. The y-axis represents how much the feature values are correlated with positive or negative image scores. As can be seen, the lower minimum lightness values are positively correlated, and the high values are negatively correlated.

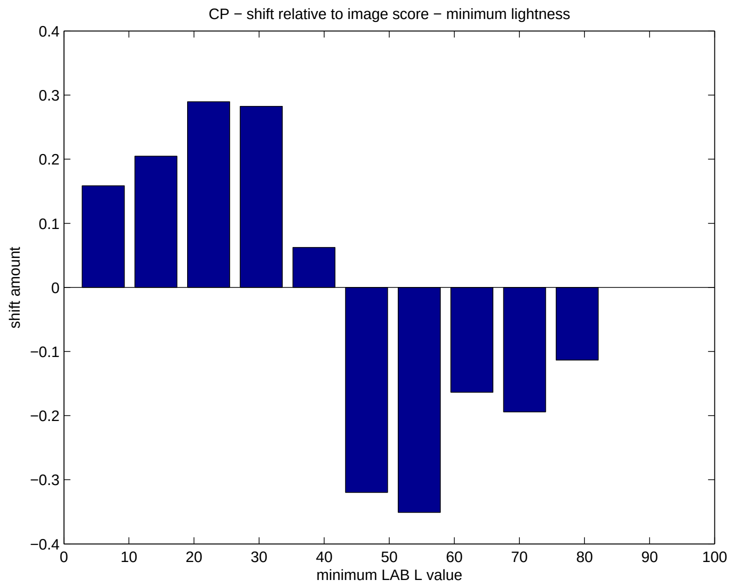


Figure 5.22: Correlation histograms for $\min(L^*)$, in *CP*

The second feature is $\max\Delta L^*$, the maximum change in lightness between two consecutive colours. A very similar feature, $R(L^*)$, the range of lightness values in an image is in third place. The correlation histogram of the *range* of the three $L^*a^*b^*$components within an image can be seen in fig. 5.23. Single colour images are omitted.

This indicates that images that do not have a large range of lightness are negatively correlated with score, whereas those with a large range of lightness, peaking at a $R(L^*)$value of 55 seem to be the most popular. As can be seen, the range of the $a^*$and b^*dimensions had comparatively little significant changes relative to image popularity.

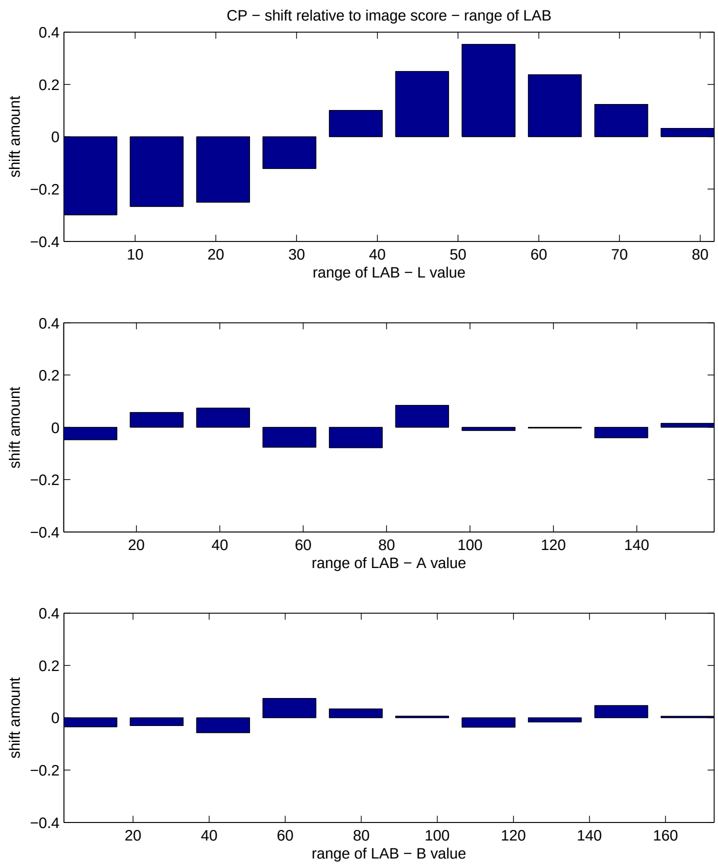


Figure 5.23: Correlation histograms for $R(L^*)$, $R(a^*)$and $R(b^*)$, in *CP*

This lends support to hypothesis number 11 – *preference for large lightness contrast* – suggested by the work of Schloss & Palmer(Schloss & Palmer, 2011).

With regards to the highly ranked $\min(L^*)$features, this also matches the large contrast hypothesis, as the lower the minimum lightness in an image, the higher the chance of their being a large range of lightness.

The next clustering of features relate to mean hue, $\mu(h^\circ)$. These are the highest ranked features representing an aspect of the mean of a colour. The correlation histogram can be seen in fig. 5.24. In this case, the single colour images have been included.

As can be observed it consists of two very significant peaks, a large positive peak at approximately 270° and a strong negative peak at about 100°. They correspond to blue and yellow respectively; blue is by far the most preferred hue, and yellow is by far the most disliked. This correlates perfectly with the studies on single colour preference, where pretty much all studies on hue preference ranked blue top, and yellow bottom. This lends strong support for hypotheses

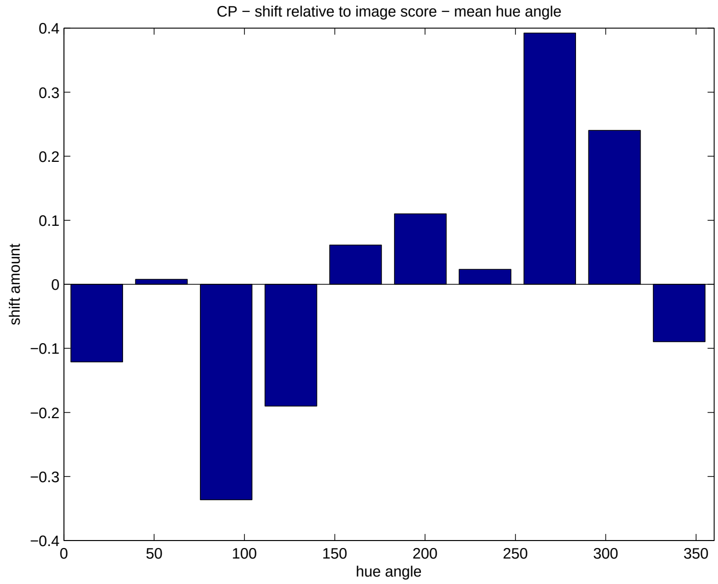


Figure 5.24: Correlation histograms for mean hue angle $\mu(h^{\circ})$, in *CP*

1 and 2 respectively, discussed in section 5.3.3.

There are two smaller peaks, one positive one at 200°, and a negative one at 0°. This corresponds to green and red respectively. Although not as significant, this seems to disagree with hypotheses 3; preference for red hues.

After hue features, another feature related to change in lightness appears, the standard deviation of lightness values, $\sigma(L^*)$, where high values are positively correlated, again supporting maximum lightness contrast.

However from there, there is quite a significant drop in *total shift* value, and so feature relevance must be interpreted more cautiously. There are a couple of features related to b^*, that correlate negatively (i.e. push towards blue, as might be expected from the results of the hue features).

The following feature, $\mu\Delta(\Delta E^*)$, the mean overall $L^*a^*b^*$distance between changing colours, has a negative correlation to small distances, and a positive correlation with high values. This means there is a preference for colours far apart in colour space. Observations of the other features $\mu\Delta$ features in the composite components sheds further light; $\mu\Delta a^*$, $\mu\Delta b^*$, $\mu\Delta C^*$, and $\mu\Delta h^{\circ}$have no discernible pattern or strong shift. This indicates the $\mu\Delta(\Delta E^*)$value is dominated by the effect of the $\mu\Delta L^*$(in fourth place); the high lightness contrast preference is stronger than any distance preference in the other colour dimensions.

The next features introduced are means and minimums of chroma, C^*. In both, the high values are negatively correlated, this indicates a dislike for high saturation images. This would initially appear to contradict hypothesis 5, *preference for high chroma colours*. However that hypothesis was with regards to single colour preference, and so perhaps not so surprising that it does not seem to apply to a Markov image.

It becomes increasingly tenuous to attribute significance the further down one goes the ranking, as the *total shift* amounts decrease. However other features of interest can be found by looking at the *absolute peak* value. By sorting the table by this value, one finds some features that have disproportionally high single peaks relative to *total shift*, this means that they will have a small range of values that have had strong correlation, but that the remaining values are not so clear cut.

The main such features relate to the information properties of the output sequences, namely the conditional entropies, $H_{Y|X}$, $H^2_{Y|X}$, and $H^3_{Y|X}$. Correlation histograms of these conditional entropy features can be seen in fig. 5.25, again single colour images were omitted, as these features do not apply. This indicates that images with very low conditional entropy, that are perfectly or nearly perfectly repetitive, were not desirable, this applied to all three orders of repetition[13]. Images of maximum entropy were neither particularly desirable or undesirable. In between, it is difficult to identify a pattern with certainty.

In summary, the following observations were made:

- Observation 1: *Dislike for single colour images*
- Observation 2: *Preference for high lightness contrast*
- Observation 3: *Preference for blues; aprox. $h°275°$*
- Observation 4: *Dislike for yellows; aprox. $h°90°$*
- Observation 5: *Preference for purples; aprox. $h°300°$*
- Observation 6: *Dislike for green; aprox. $h°130°$*
- Observation 7: *Dislike for colours of high chroma (highly saturated colours)*
- Observation 8: *Dislike for perfect repetition*

[13]As a reminder, a first order zero entropy sequence could be *A B A B A B*, or *A B C A B C*, a second order zero entropy sequence could be *A B B A B B A B B* or *A B B C A B B C A B B C* and so on

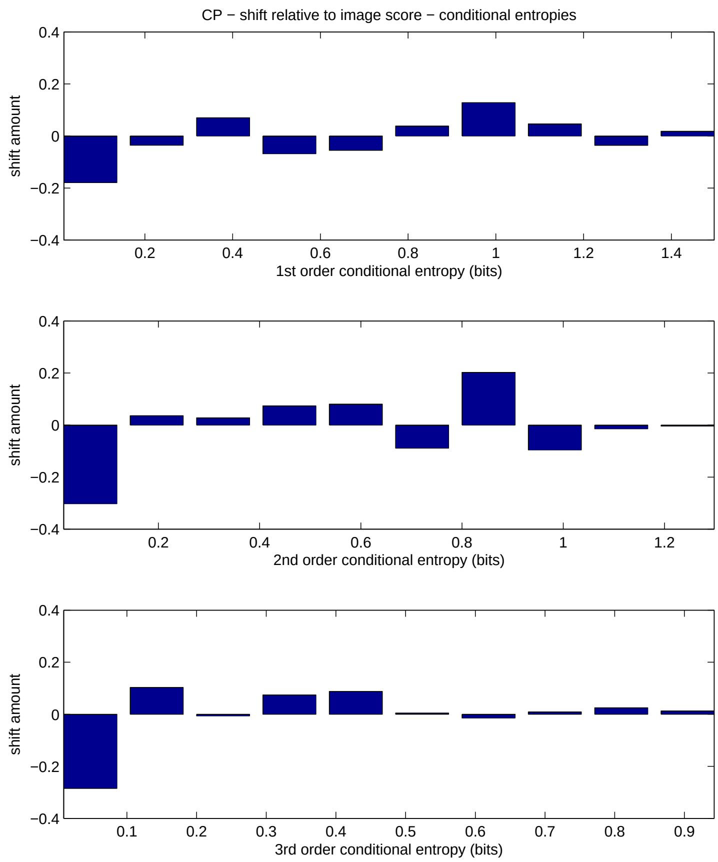


Figure 5.25: Correlation histograms for conditional entropies in *CP*; $H_{Y|X}$, $H^2_{Y|X}$ and $H^3_{Y|X}$

The next sections consider the analysis of the evolving population.

5.6.2 Measuring Fitness Change in Evolving Populations

As mentioned in the 5.4.2, once every four selections user would do a 'population comparison', comparing the fitness of an image from the newest generation of a population verses that of the first. As the number of such comparisons before a new generation is quite small (on average only consists of 30 selections[14]), it is expected that there would be a large amount of noise and variation in these comparisons. Nonetheless it is possible to gleam some patterns in the fitness of the populations, as illustrated in fig. 5.26.

As can be seen, $EP3$ was the only evolution that had a clear fitness improvement over its starting generation. The $EP1$ and $EP2$ plots are noisy and not indicative of any fitness improvement. Given that, as discussed previously, $EP1$ and $EP2$ evolved towards images of maximum entropy, yet were still evolving their colour pallets, the following conclusion could be drawn.

- Observation 9: *Dislike for maximum entropy*

To establish if there were any significant difference in overall aesthetic value between $EP1$ and $EP2$, once the evolutions were suspended and $EP3$ was launched, population comparisons for $EP1$ and $EP2$ were re-established and left to measure the difference in fitness over a larger number of comparisons. These ran over 336 and 340 selections respectively, and as can be seen in table 5.10, EP2 fared significantly better, establishing a score of +136 whereas of $EP1$ only had a score of +26.

	EP1 Gen70 vs Gen1	**EP2 Gen70 vs Gen1**
Wins	150	208
Losses	124	72
Draws	62	60
Count	336	340
Score	**26**	**136**

Table 5.10: Population comparisons for $EP1$ and $EP2$

This will be down to the only differing variable between the two, the 'elitism' of $EP2$, whereby selected parent images would stay in the population. This determines that 'elitism'

[14] One might expect that doing a 'population comparison' selection once every fourth selection, given each generation is 100 selections, that the average number would come to 25. The reason the average is about 30 relates to an implementation detail. The generation of a new generation and the images mating process would run in a background thread so as to not have the users wait for the process to complete. During this time, the evolving population would be but 'on hold', and no selections would be made on them. Instead the users would be doing only 'population comparisons' until the next generation was ready.

to some extent counteracts biases in the evolution mating processes. It must be noted that there may be other, undetected, biases in the evolutionary processes beyond those of the tendency towards high entropy of $EP1$ and $EP2$. There is no guarantee that the implementation in $EP3$ is by any means 'perfect' and free of biases. However the measurable increase in fitness reassures that the clusterings of features in the higher generations should to some extent correlate to preference. $EP1$ had just too low a fitness improvement however, and chances are that any clusterings found within would be misleading. As such only $EP2$ and $EP3$ will be analysed, and not $EP1$.

The difference between the images of $EP2$ and $EP3$, is that $EP2$ images have *on average* Markov chains of higher entropy. This does not necessarily mean that all images in $EP2$ are of high entropy; the elitism can ensure that popular low entropy images stay in the population for a long time, and further the 'seed values' that form part of the genome can influence how the resulting image appears. On the whole however, $EP2$ images are significantly busier than $EP3$. Comparing how differently colour distributions end up being between $EP2$ and $EP3$ can indicate in what respect colour preference are contingent on the entropy of the image.

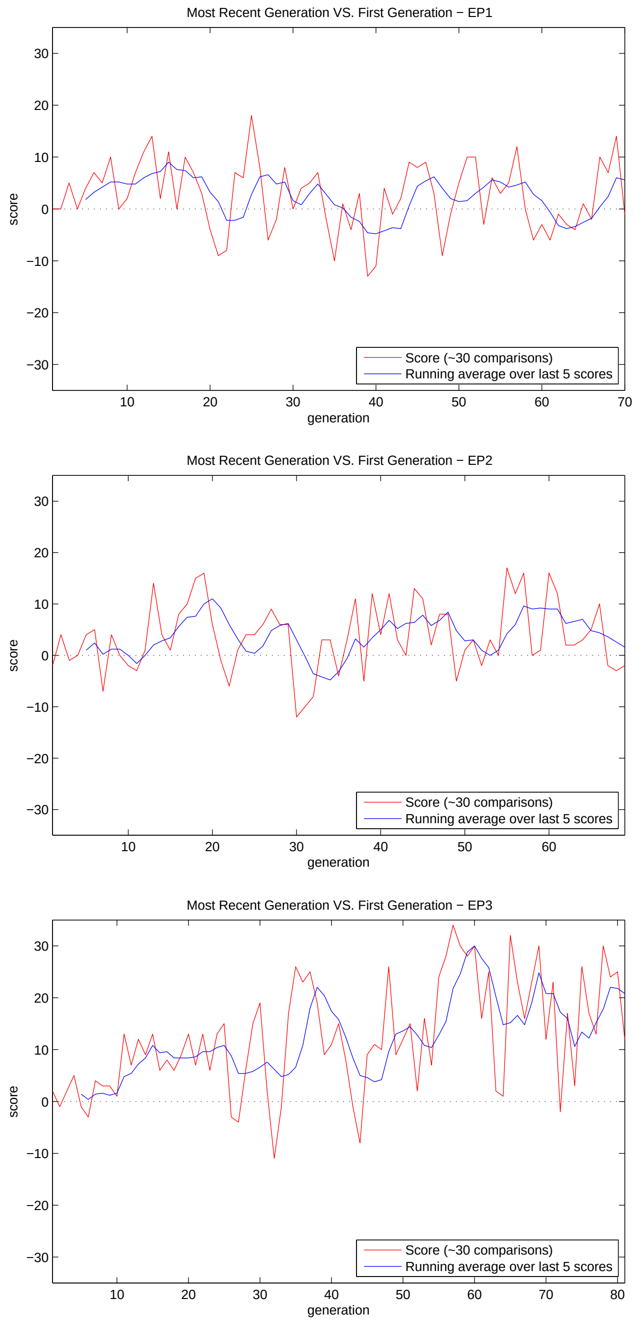


Figure 5.26: Population comparisons results for $EP1$, $EP2$and $EP3$. The red lines are the actual values, and consisted of on average 30 selections. The blue line is the running average of the last five actual values.

5.6.3 Analysis of EP2

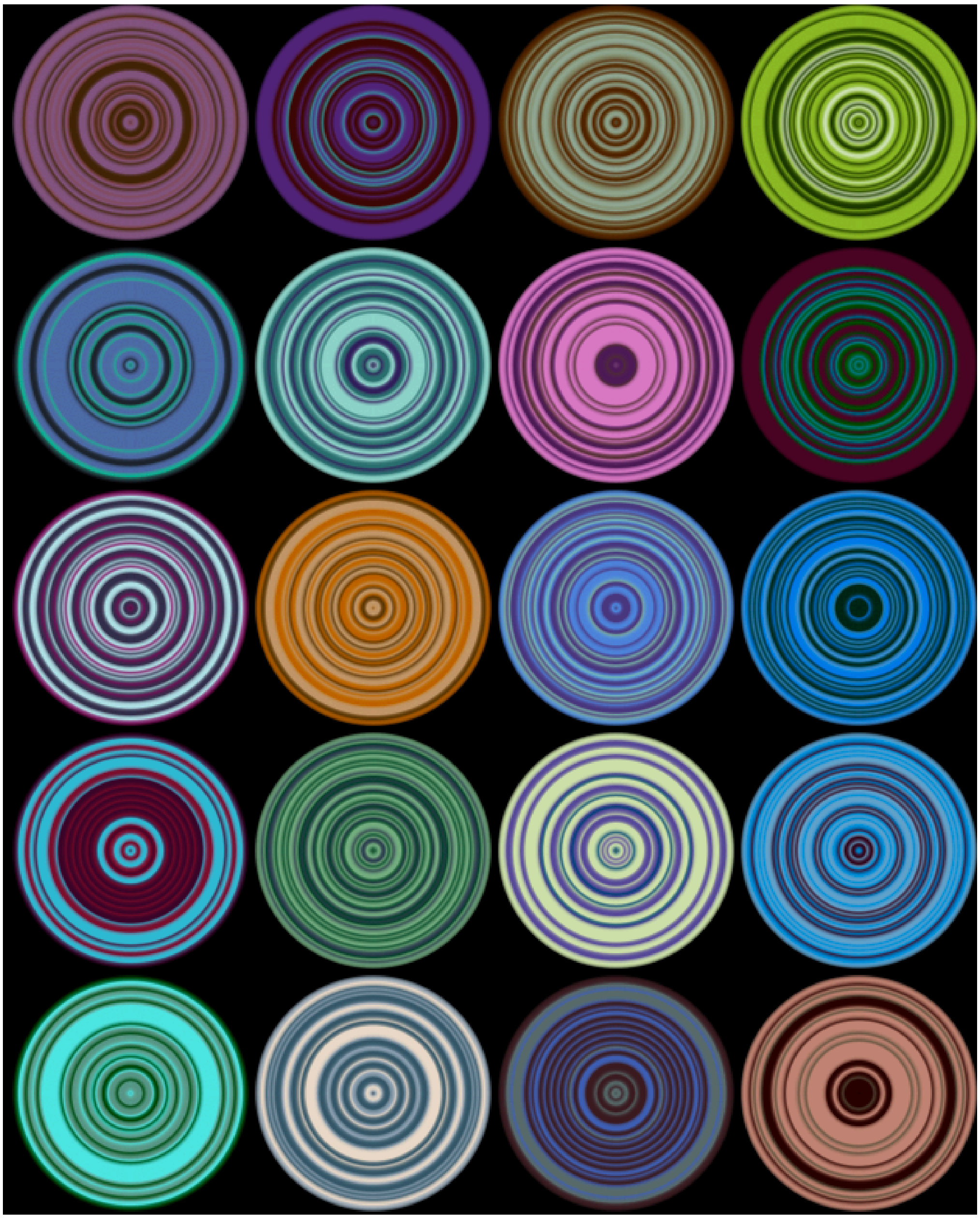


Figure 5.27: 20 images from generation 70 of $EP2$.

Fig. 5.27 and fig. 5.28 show 40 images randomly selected from the final generation of $EP2$. It is difficult to observe overall trends by eye, there is no obviously dominant hue. As can be seen many images are quite busy, that is, of high entropy, but that is not exclusively the case. All the images of the final generation of $EP2$ are available as item $Ex5$ of the illustrative materials accompanying this thesis.

Figure 5.28: 20 more images from generation 70 of $EP2$.

The first step taken in the analysis was to attempt to identify the features that have had the biggest overall changes in value, as well as identifying the features that tended to become more homogenised as the generations progressed. Table 5.11 summarises the feature changes between the first and last generation. Only features who's values are not influenced by the entropy bias are presented; a shorter list than previously. Further, single colour images were omitted when calculating the values from this table, as it is known that they are unpopular and would vanish due to the entropy bias, and their feature values which are often zero would make it harder to identify trends.

Table 5.11: Features of images in $EP2$. $\mu gen1$ is the mean of the values in generation 1, $\mu gen70$ is the mean of the values in generation 70, $\Delta\mu$ is the change between the two means. $\sigma gen1$ is the standard deviation of generation1, $\sigma gen70$ is the standard deviation at the final generation. $\Delta\sigma$ represents the *homogenisation*, which is the relative magnitude of the decrease in standard deviation between the fist and the last generation. Table is sorted by decreasing *homogenisation*.

Feature	$\mu gen1$	$\mu gen70$	$\Delta\mu$	$\sigma gen1$	$\sigma gen70$	$\Delta\sigma$
$\min\Delta b^*$	26.49	12.75	-13.74	24.64	10.81	0.56
$\min\Delta a^*$	25.14	12.14	-13.00	22.19	10.31	0.54
$\min\Delta C^*$	15.67	8.43	-7.25	14.08	7.50	0.47
$\min\Delta h^\circ$	47.93	30.76	-17.17	40.48	24.70	0.39
$\mu(a^*)$	6.20	3.35	-2.85	28.16	17.27	0.39
$\min\Delta(\Delta E^*)$	29.86	25.33	-4.53	14.99	9.48	0.37
$\mu\Delta a^*$	48.17	29.85	-18.33	25.98	16.68	0.36
$\mu(b^*)$	4.62	-11.05	-15.67	30.85	20.32	0.34
$\mu\Delta b^*$	51.23	32.78	-18.45	29.50	19.83	0.33
$\max\Delta(\Delta E^*)$	64.03	65.70	1.67	20.98	14.40	0.31
$\min(C^*)$	36.05	22.33	-13.72	17.26	11.96	0.31
$\mu\Delta(\Delta E^*)$	47.76	47.04	-0.71	15.92	11.11	0.30
$\min(a^*)$	-29.41	-19.14	10.27	29.72	20.75	0.30
$\min(b^*)$	-33.03	-34.91	-1.88	36.02	25.69	0.29
$\max(a^*)$	40.03	25.33	-14.71	29.10	20.88	0.28
$\max\Delta a^*$	68.25	44.46	-23.79	33.29	24.03	0.28
$R(a^*)$	69.44	44.46	-24.98	33.17	24.03	0.28
$\mu(C^*)$	55.89	37.77	-18.12	17.33	12.64	0.27
$R(b^*)$	72.61	49.11	-23.50	38.27	28.28	0.26
$\max\Delta b^*$	71.78	49.11	-22.68	38.18	28.28	0.26
$\min\Delta L^*$	11.95	13.64	1.70	11.35	8.86	0.22
$\mu\Delta L^*$	23.40	34.13	10.73	14.44	11.41	0.21
$\max(b^*)$	39.58	14.20	-25.38	30.59	24.44	0.20
$\mu\Delta C^*$	28.40	21.41	-6.99	16.70	13.36	0.20
$\max\Delta L^*$	32.74	51.07	18.34	18.70	15.25	0.18
$R(L^*)$	33.13	51.07	17.94	18.69	15.25	0.18
$\max(h^\circ)$	272.15	286.39	14.24	73.93	61.09	0.17
$\mu\Delta h^\circ$	90.08	77.98	-12.11	40.84	34.89	0.15
$\min(L^*)$	40.84	25.50	-15.34	15.67	13.61	0.13
$\mu(h^\circ)$	173.39	228.45	55.06	109.35	95.18	0.13
$R(C^*)$	41.38	32.62	-8.77	21.96	19.84	0.10
$\max\Delta C^*$	40.73	32.62	-8.11	21.80	19.84	0.09
$\mu(L^*)$	57.51	50.53	-6.98	13.85	12.74	0.08
$\max\Delta h^\circ$	121.94	113.06	-8.87	48.34	46.39	0.04
$\max(L^*)$	73.97	76.58	2.60	13.67	13.57	0.01
$\max(C^*)$	77.43	54.95	-22.49	19.56	19.43	0.01

As can be observed, the greatest homogenisation occurred with respect to the *minimum difference* between adjacent colours in an image. The overall trend in the $\min\Delta a^*$, $\min\Delta b^*$, $\min\Delta C^*$and $\min\Delta h^\circ$features favour a decreases in the minimum difference. The other features that relate to the difference between colours, such as the $\mu\Delta$ of these components also reinforced a decrease in these values. However that did not apply to the lightness component, as can be seen the $\mu\Delta L^*$value increased.

The overall trend of the difference between adjacent colours can be observed in fig. 5.29, where correlation histograms representing the difference between the first and last generation, of the mean difference, $\mu\Delta$, in adjacent colours is represented.

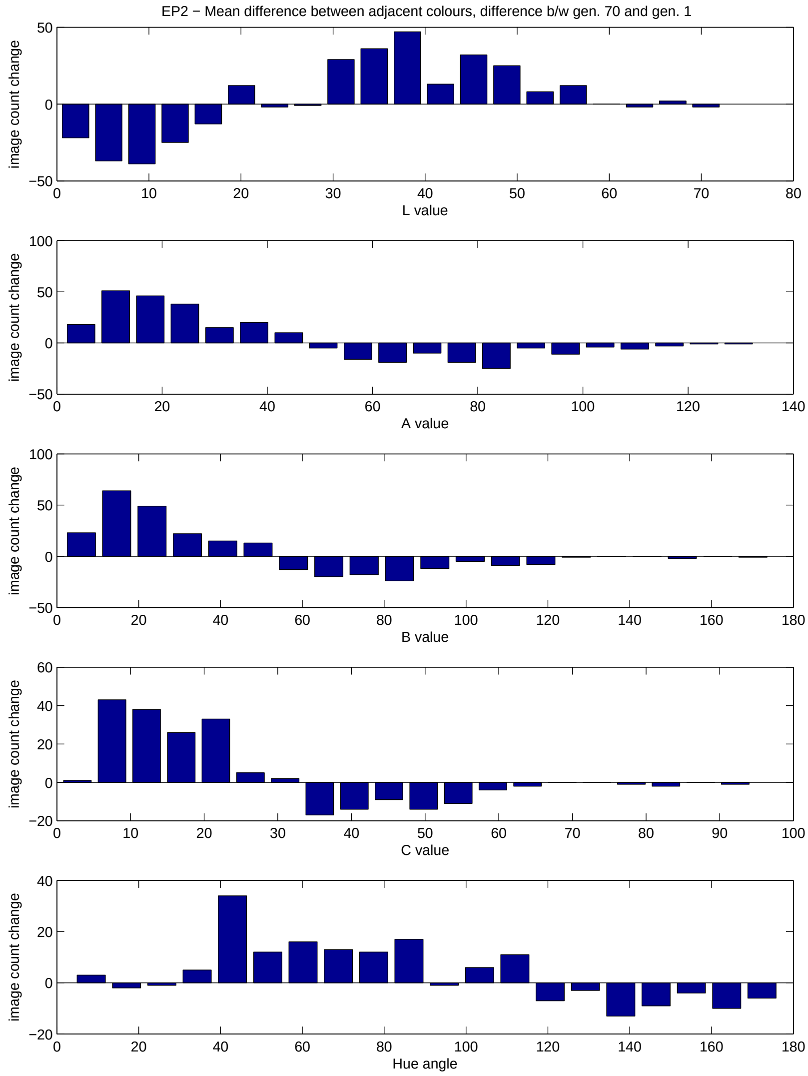


Figure 5.29: Histograms comparing the first and last generations of $EP2$ for the $\mu\Delta L^*$, $\mu\Delta a^*$, $\mu\Delta b^*$, $\mu\Delta C^*$, and $\mu\Delta h^\circ$features, representing the mean change in adjacent colours. (Note these histograms may not 'even out', this due to single colour images not having defined value for these features)

As can seen the colour difference of adjacent colours with regards to the L^*decreased for low values, but increased for higher values. This reinforced *Observation 2*, that high lightness contrast images are favoured. However with respect to the a^*, $b^*$and C^*components, the opposite seems trues, and the low differences are reinforced. It is then possible to add the following observations:

- Observation 10: *Preference for images with similar a^* values*
- Observation 11: *Preference for images with similar b^* values*
- Observation 12: *Preference for images with similar C^* values*

Some care must be taken in interpreting the negative correlation values when observing such histograms comparing the generations of an evolution. Due to the nature of the darwinian evolutionary process, the *most* preferred attributes will be reinforced the strongest. If a feature attribute is quite favoured, but not as much as an another attribute that is favoured even more, it could end up having a negative value. This depends on how homogenised the populations become. In other words, there could well be a 'local' maxima of preference at a feature value that is negatively correlated. Therefore attributing a 'dislike', or rejecting a hypothesis, due to a negative correlation when comparing populations of an evolution must be done cautiously.

At first glance it would appear that *observation 10* and *observation 11* support *hypothesis 8*; preference for similar hues. However a closer reading of the hue plot indicates that this is a little more nuanced. Hue differences above 120°appear to have decreased in the population, but the negative correlations at close to 180°(representing *hypothesis 9* - opponent hues), is of low magnitude, as is the correlation for 120°(representing *hypothesis 10* - triadic harmony[15]). This is not evidence enough to reject them, but there is no evidence to support them either. With regard to supporting the preference for similar hues hypothesis, there seems to be on the whole a trend towards that, but colour combinations less than 40°apart do not seem to necessarily have a preference. However the following observation can be made:

- Observation 13: *Preference for images with colours 40°to 90°apart in hue*

Strong homogenisation also occurred with respect to mean values of the $b^*$and a^*components. The $\mu(b^*)$had a significant shift towards negative values, that is towards blue and away from yellow. This supports previous observations of blueness being preferred over yellow. Given the weight of evidence in the literature for this, and further the evidence of this trend in *CP*, this demonstrates that, with respect to this feature at least, the direction of evolution seems to correlate with preference.

[15]Due to the entropy bias of $EP2$, nearly all images had three colours.

The a^*component also homogenised strongly, but did not shift the mean significantly. Instead the outer extremes of the a^*component decreased, as can be seen by the increase in $\min(a^*)$and decrease in $\max(a^*)$.

With regards to lightness, the mean, $\mu(L^*)$, and $\min(L^*)$decreased significantly, yet did not homogenise significantly. Conversely however it interesting to note that the maximum change in lightness, $\max \Delta L^*$, homogenised strongly and had a slight increase, and the range of lightness, $R(L^*)$, increased significantly. This seem to indicate that although somewhat more somber images were preferred, a high lightness contrast was reinforced. This is confirmed in fig. 5.30 where the difference in the range of lightness is shown between the first and last generation.

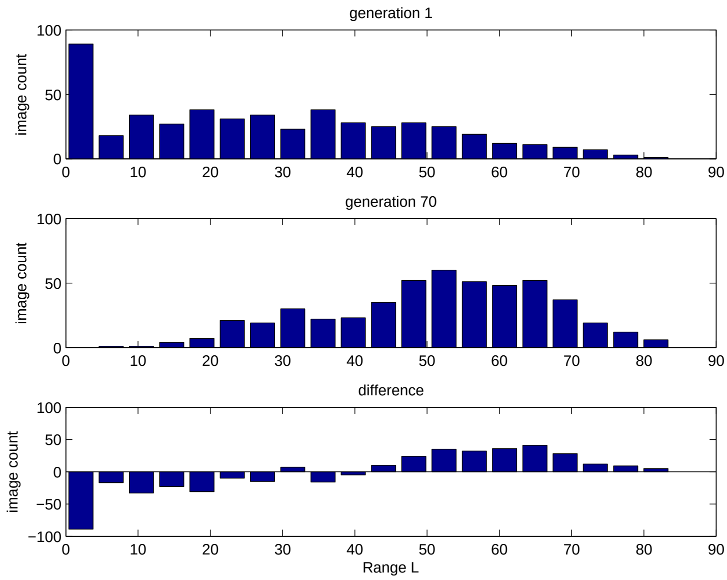


Figure 5.30: Histograms for $R(L^*)$comparing the first and last generations of $EP2$.

As can clearly be seen, range of zero, i.e. single colour images, have vanished completely by the last generation. The remainder of the histograms match *Observation 2*, that high lightness contrast images are favoured. This indicates that the evolution was also able to carry across this aspect of user preference.

To get a sense of how the overall mean colour shifted from generation to generation, their $L^*a^*b^*$components are shown, together with renderings of the mean colour, in fig. 5.31. A clearer understanding of the overall colour properties of $EP2$ can be had by considering fig. 5.32. These plots were generated by counting the visible colours in all the images in generation 70, and calculating their distributions in a histograms for each of the $L^*a^*b^*$and $L^*C^*h^\circ$components. These

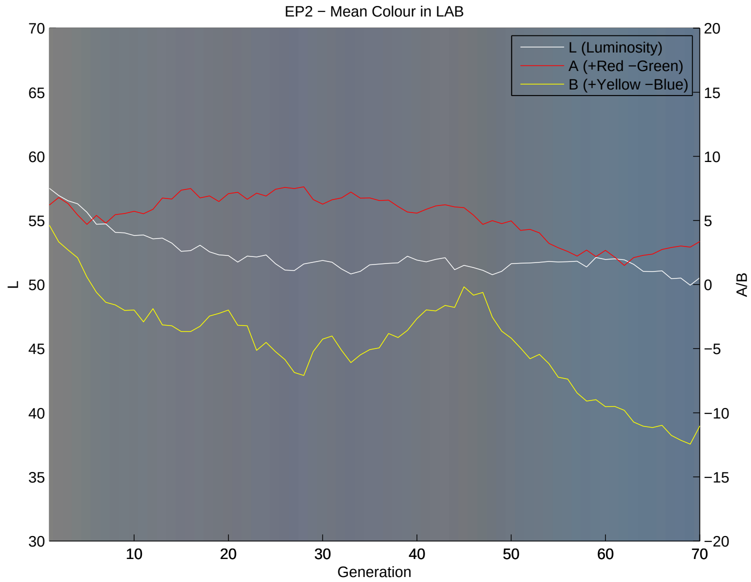


Figure 5.31: Mean values of L^*, a^*, and $b^*$for each generation in $EP2$

were then normalised, and subtracted from the distributions of the randomly generated sRGB values presented earlier in fig. 5.8. This then results in a correlation histogram indicating what values of these colour dimensions were reinforced or suppressed by the evolutionary process.

These correlation histograms show clearly how the darker colours, of L^*less than 40 were favoured. This observation can be added.

- Observation 14: *Preference for images with mean low lightness*

White and near white was also reinforced (L^*values 90 and above), however perhaps not strongly enough to be sure of statistical significance.

It is also easy to see the homogenisation of a^*, as well as the shift towards in the b^*towards the negative. Interestingly the most extreme negative values did not have an increase. This may be related to what happened with regards to chroma, C^*. As can be seen, the there is a clear preference for low chroma values, particularly below 40, and clear dislike for increasing saturation, 60 and above. This also was seen in the analysis of *CP* with *Observation 7*.

With regards to hue, $h°$, there is a dislike for the whole range from about 30°, red/orange, via (mild dislike for) orange and yellow, all the way to 150°(green with slight yellow). There is particular dislike for angles of about 140. This corresponds to yellow/green. It has already been observed that there was a strong dislike for yellow and green in *CP*, with *Observation 4* and *Observation 5*.

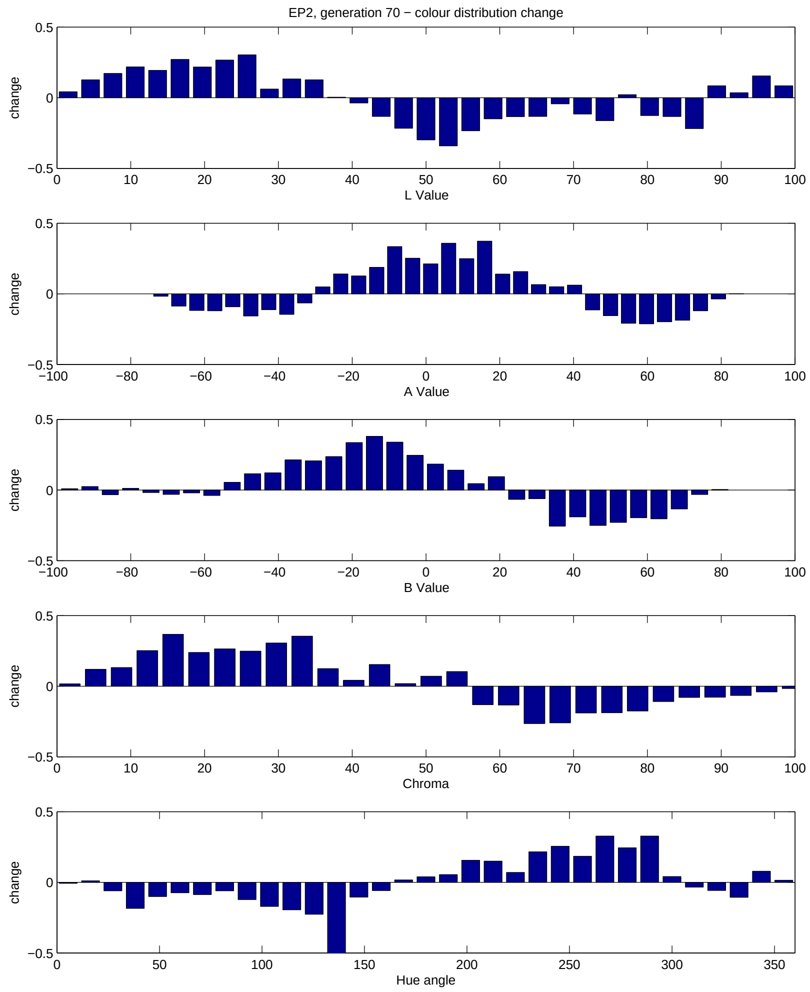


Figure 5.32: Correlation histograms for the final generation of $EP2$ for all colour dimensions in $L^*a^*b^*$and $L^*C^*h^\circ$

Conversely there is a preference for hues ranging from about 200°(green/blue) via blue, all the way to 290°(purple). This as already been established in *CP*, with *Observation 3* and *Observation 5*.

This view however gives broader reading of the preferences of hues that was acquired in *CP*. It is thus worth drawing these observations.

- Observation 15: *Preference for h°range - 200°to 300°*
- Observation 16: *Dislike for h°range - 30°to 150°*

The range from 300°to 30°are largely red, and have no preference or dislike.

Many of the observations derived from the analysis of *CP* were observed again in the analysis of *EP*2; this includes the high lightness contrast preference, the blue over yellow preference, and even the preference for low chroma. This is evidence that the direction of the evolution of overall colour pallets correlates well with user preference.

5.6.4 Analysis of EP3

Figure 5.33: 20 images from generation 82 of *EP*3.

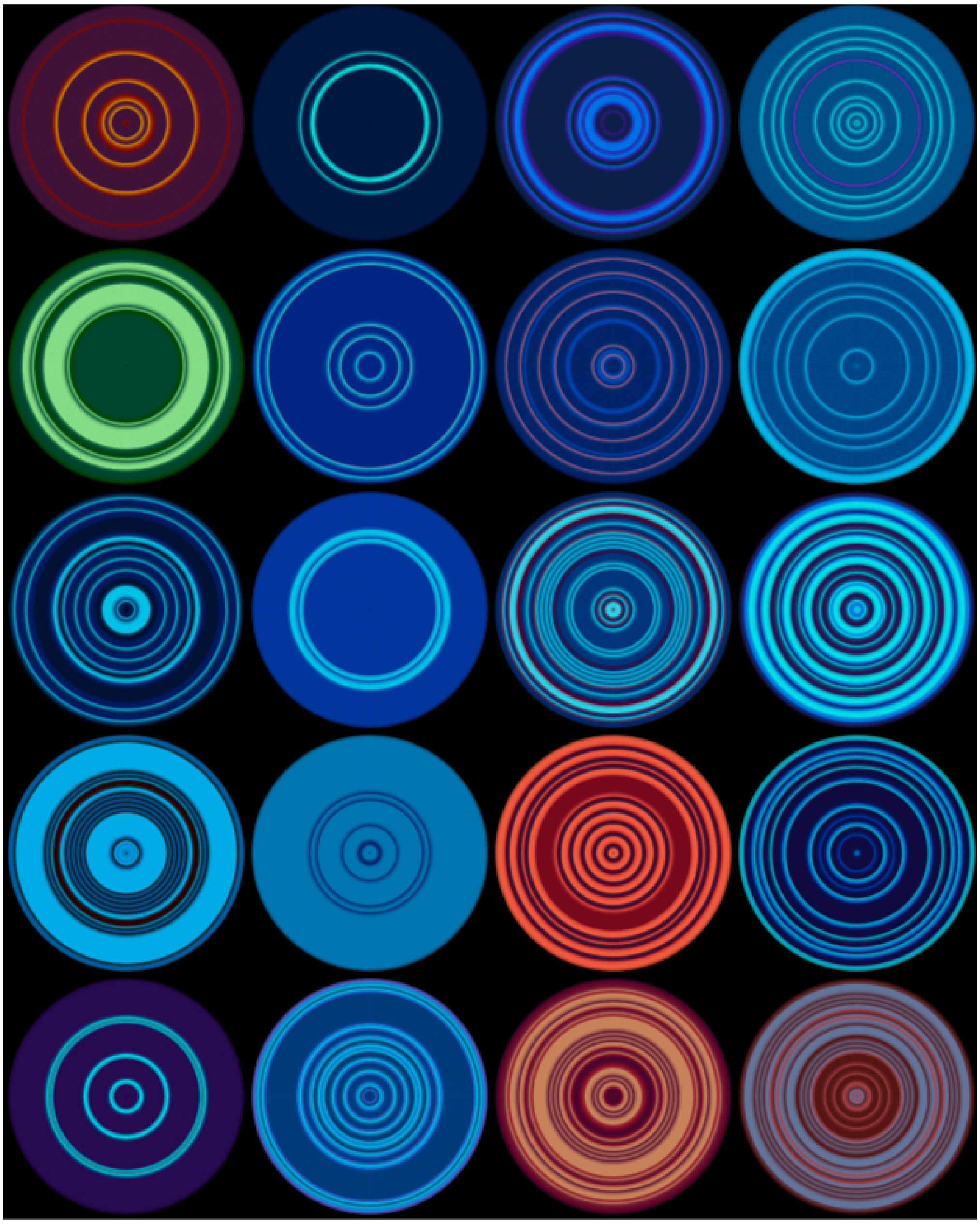


Figure 5.34: 20 more images from generation 82 of $EP3$.

Fig. 5.33 and fig. 5.34 show forty images randomly selected from the final generation of $EP3$. From visual observation, there still seems to be considerable variety in their visual attributes. They seem to be less visually busy than the images from $EP2$, with many being quite sparse. There also seems to be a greater proportion of blue images. To best be able to identify how these images differ from those in $EP2$, the analysis will begin by observing the attributes of the images considered in the previous section, the overall colour attributes and the mean change between neighbouring colours. All the images of the final generation of $EP3$ are available as item $Ex6$ of the illustrative materials accompanying this thesis.

Fig. 5.35 shows how the mean colours of $EP3$ changed with each generation.

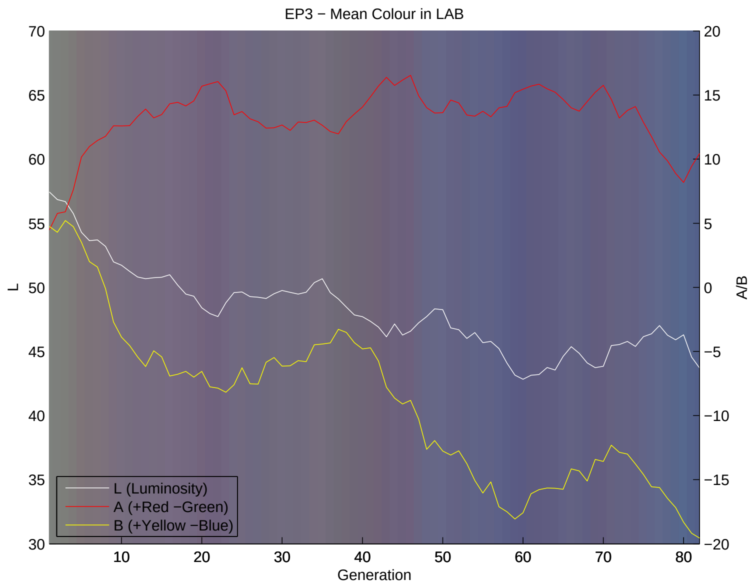


Figure 5.35: The mean values of L^*, a^*, and $b^*$for each generation in $EP3$

Comparing fig. 5.35 with fig. 5.31, it can be seen that the trajectory of the mean L^*components have some differences. Considering only the first 70 generations of $EP3$ to match the 70 of $EP2$, for mean lightness, $\mu(L^*)$, is at about 44 for $EP3$ and at 50 for $EP2$. It could be conjectured that for lower entropy images, low lightness is preferred, however the difference is fairly low. Similarly the b^*component has decreased more by generation 70, but by not enough to attribute strong significance to.

What is more striking is the differing trajectories of the a^*component; whereas in $EP3$ it initially increases significantly and stays high up for a long time, in $EP2$ it effectively stays the same level the whole time. At generation 70, the difference between the two a^*values is 15, a considerable amount. As $EP3$ contains higher entropy images than in $EP2$, this leads to the following observation:

- Observation 17: *Preference for high a^*values is greater in low entropy images*

A difference in $L^*a^*b^*$components indicates a difference in either the hue or chroma (or both). Comparing the progression of mean hue yielded no major differences between $EP2$ and $EP3$, however an interesting difference is evident by plotting the mean chroma, $\mu(C^*)$, of $EP2$ and $EP3$. This is shown in fig. 5.36.

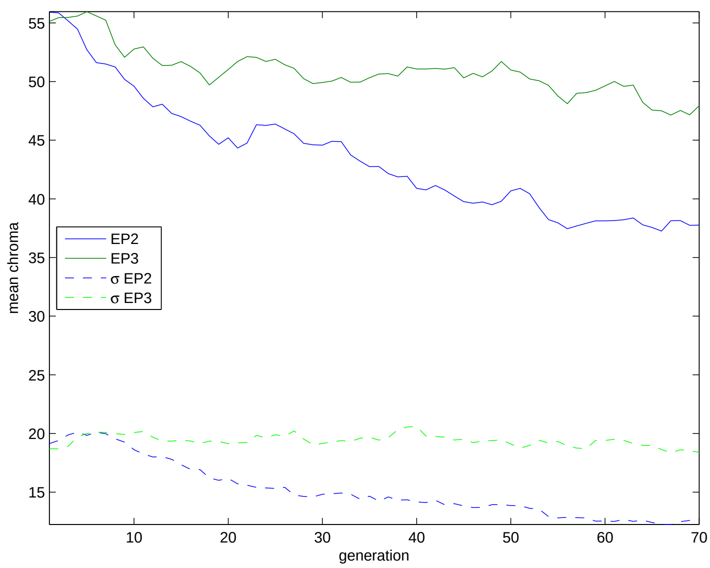


Figure 5.36: The mean chroma, $\mu(C^*)$, and their standard deviations, for $EP2$ and the 70 first generations of $EP3$

Clearly, in $EP2$ there is a significantly stronger decrease and homogenisation than in $EP3$. That is, if colours change frequently, it is preferred to not have high saturation in the colours. From this the following observation can be made:

- Observation 18: *Preference for low C^*values is greater in low entropy images*

The overall colour distributions of $EP3$ can be seen in fig. 5.37.

On an overall glance, there are some very similar properties as there are in fig. 5.32, the corresponding plot for $EP2$. For all colour components, the same overall shape is evident. There are differences, usually the $EP3$ plot is more jagged and noisier, particularly for chroma. As was noted already, there was less homogenisation in that colour dimension in $EP3$.

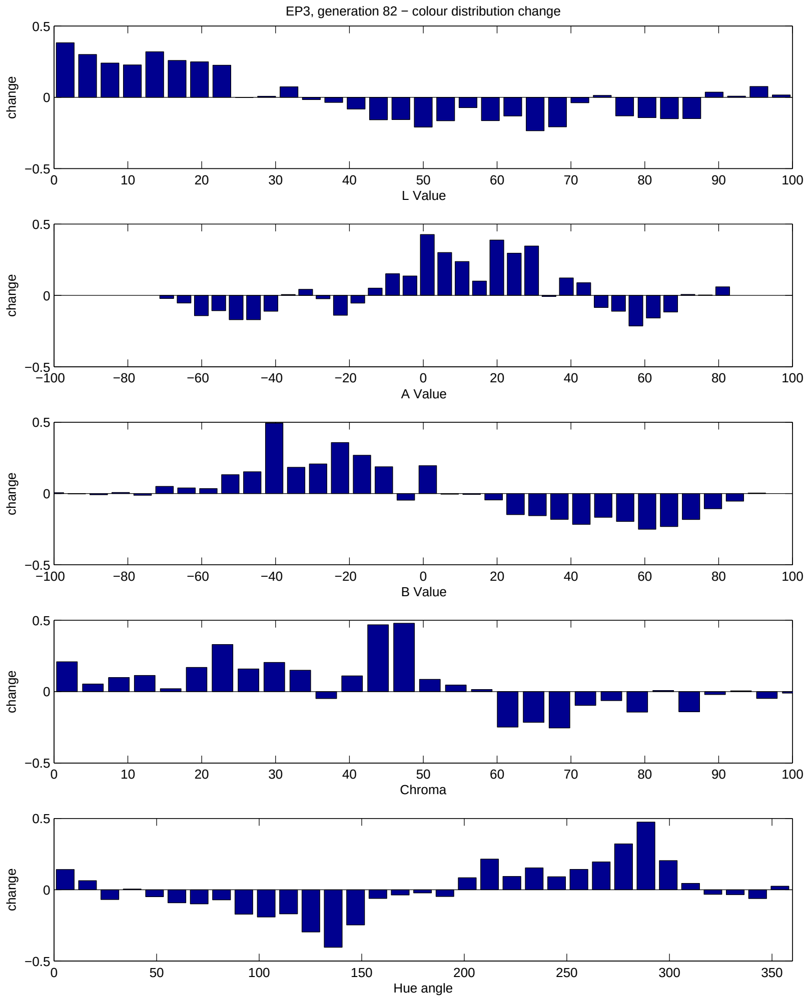


Figure 5.37: Correlation histograms for the final generation of $EP3$ for all the colour dimensions in $L^*a^*b^*$and $L^*C^*h^\circ$

There is an interesting difference between the distribution of the L^*values. In $EP3$ there is a significantly stronger preference for images of minimum or near minimum lightness, that is black or near black. From this the following observation can be made:

- Observation 19: *Preference for black and near black is greater in images of low entropy*

Characteristics of the *change* in adjacent colours of $EP3$ can be seen in fig. 5.38. Again there is largely the same shapes as in the corresponding plot for $EP2$, in fig. 5.29. The main difference is in the chroma, $\mu\Delta C^*$, which is unsurprising with regards to the already mentioned differences of that component between $EP2$ and $EP3$. There is a notable difference however in the preference in hue between adjacent colours, $\mu\Delta h^\circ$. In $EP3$ there is a clear preference for differences in hue angle below 40°. In fact, that region is the area with the strongest preference, lending support for *hypothesis 8*: preference for similar hues. From this the following observation can be made:

- Observation 20: *Preference for similar hues in images not of high entropy.*

Before moving on from colour to discuss the information theoretic properties of the evolution of $EP3$, a visualisation of the distributions of all the colours in $EP2$ and $EP3$ in $L^*a^*b^*$space is presented. This is shown in fig. 5.39

Some of the overall trends described previously can be identified if compared with fig. 5.6, those of the random distribution. However there is also another noteworthy trend: some colours line up right at the edge of the sRGB gamut, as can be seen by the straight lines of colours, particularly in $EP3$. The reason a colour might enter the genome right on the edge is due to the colour mating algorithm described in 5.4.3, where the translation of the geometry of the colours could push colours to the edge of the sRGB gamut. However if those colours were not liked or favoured, they would disappear due to natural selection. Instead they stay in the population. The following observation can thus be made:

- Observation 21: *Some colours on the edge of sRGB gamut are preferred.*

This is perhaps not so surprising as the limitations of sRGB (and by extension most computer monitors) are widely recognised; more colours that can be seen in real life than can be rendered on screen. Modern digital cameras often provide another colour space, AdobeRGB with a larger gamut then sRGB. This allows photographers to capture the more vivid colours that can be printed, but not displayed on monitors. One can conjecture that the perhaps if those colours could be rendered that they might be preferred.

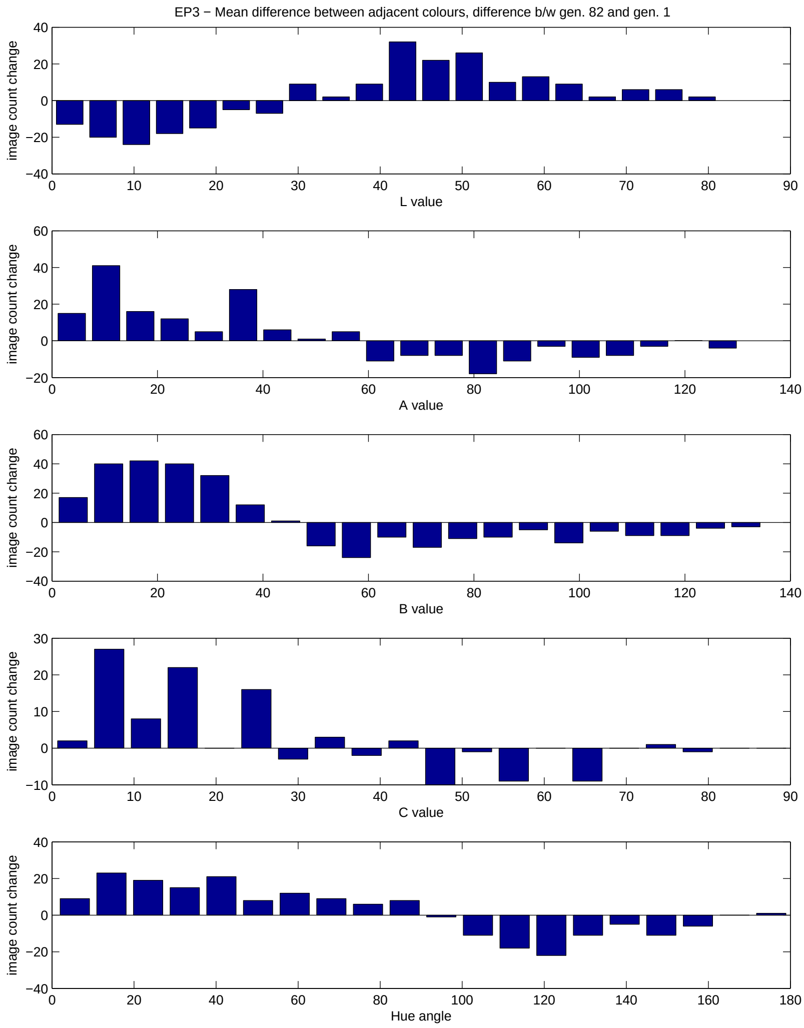


Figure 5.38: Histograms comparing the first and last generations of $EP3$ for the $\mu\Delta L^*$, $\mu\Delta a^*$, $\mu\Delta b^*$, $\mu\Delta C^*$, and $\mu\Delta h^\circ$features, representing the mean change in adjacent colours. (Note these histograms may not 'even out', this due to single colour images not having defined value for these features)

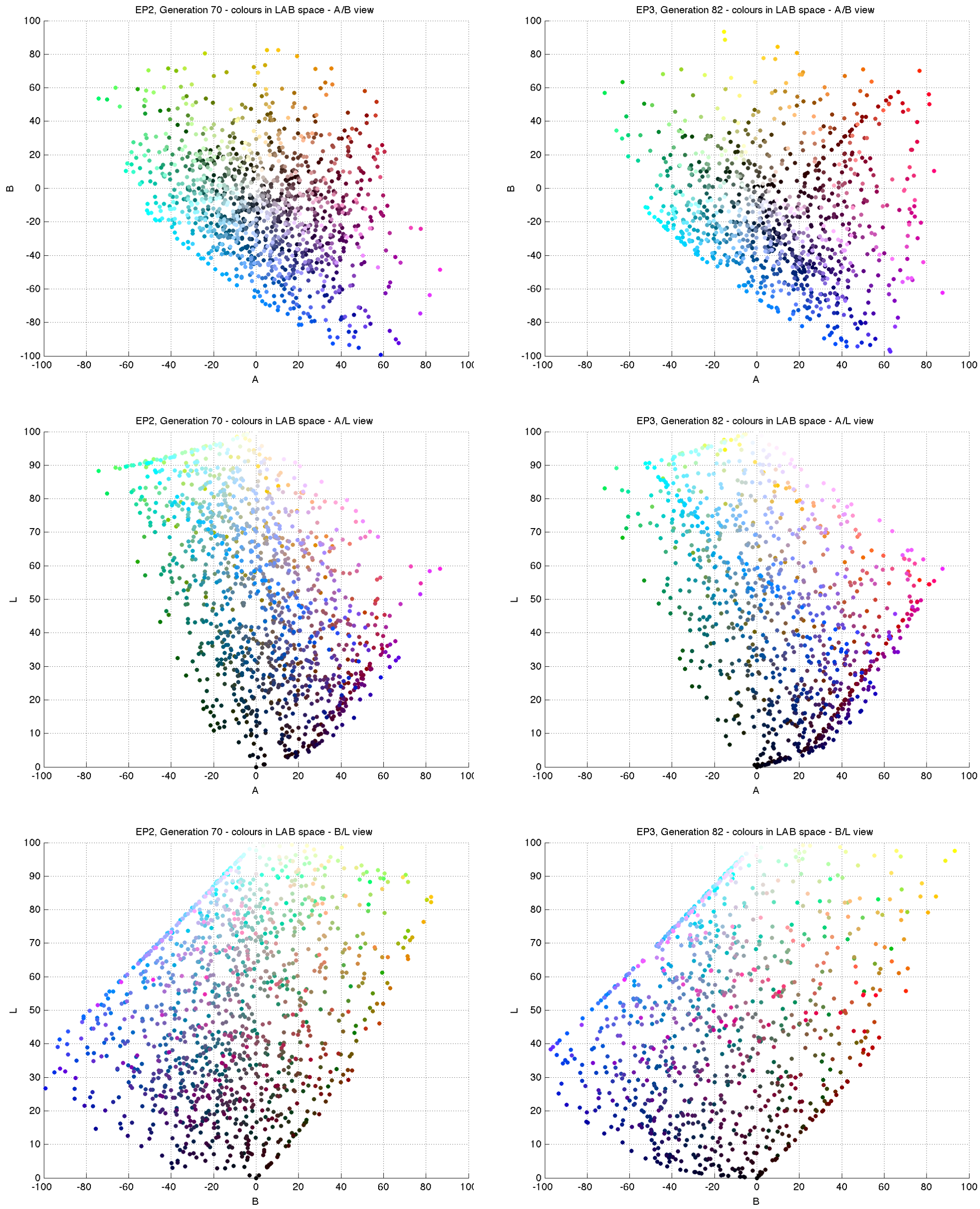


Figure 5.39: Distributions of colours in generation 70 of $EP2$ (column 1), and generation 82 of $EP3$ (column 2), in a^*/b^*view (row 1), a^*/L^*view (row 2) and L^*/b^*(row 3).

Information Measures

A complete list of the feature values comparing the first and last generation of $EP3$, is presented in table A.1, in the appendix. As aspects of colour have already been discussed at length, a version of the table with only the information related features is presented here in table 5.12.

Table 5.12: Non-colour features of images in $EP3$. $\mu gen1$ is the mean of the values in generation 1, $\mu gen82$ is the mean of the values in generation 82, $\Delta\mu$ is the change between the two means. $\sigma gen1$ is the standard deviation of generation1, $\sigma gen82$ is the standard deviation at the final generation. $\Delta\sigma$ represents the *homogenisation*, which is the relative magnitude of the decrease in standard deviation between the fist and the last generation. Table is sorted by decreasing *homogenisation*.

Feature	$\mu gen1$	$\mu gen82$	$\Delta\mu$	$\sigma gen1$	$\sigma gen82$	$\Delta\sigma$
$I_{Y;X}$	0.30	0.17	-0.12	0.27	0.16	0.43
$\max\lvert\vec{c}\rvert$	9.17	12.47	3.30	11.81	6.92	0.41
$\lvert\Delta c\rvert$	27.84	17.60	-10.24	12.85	8.50	0.34
$\lvert C\rvert$	2.82	2.92	0.10	0.39	0.27	0.31
$H^2_{Y\lvert X}$	0.63	0.67	0.04	0.31	0.22	0.28
$H^3_{Y\lvert X}$	0.44	0.52	0.08	0.22	0.16	0.28
$I^2_{Y;X}$	0.56	0.36	-0.20	0.31	0.23	0.26
$H_{Y\lvert X}$	0.90	0.86	-0.04	0.41	0.35	0.15
H_X	1.20	1.04	-0.17	0.37	0.36	0.03
$I^3_{Y;X}$	0.75	0.51	-0.23	0.36	0.35	0.01

As can be observed the greatest homogenisation occurs with respect to the *first order mutual information*, $I_{Y;X}$, and the longest unbroken run of single colour, $\max|\vec{c}|$. (These are also the top homogenising features in the complete table in the appendix). Followed by $|\Delta c|$, the number of colour changes in an image, and the number of different colours in the image, $|C|$.

The number of colours in the image increases, as can be seen in fig. 5.40, three colour images have had by far the greatest increase. The following observation can be added.

- Observation 22: *Three colour images are preferred.*

With regards to the information measures, on the whole, the mutual information measures, $I_{Y;X}$, $I^2_{Y;X}$and $I^3_{Y;X}$decrease in value. One might thus expect the conditional entropies, $H_{Y|X}$, $H^2_{Y|X}$and $H^3_{Y|X}$, to go in the opposite direction and increase in value, however interestingly that is not really the case; $H_{Y|X}$decreases, but $H^2_{Y|X}$stays roughly the same while $H^3_{Y|X}$increases.

To get a sense of what is happening, the images were placed in triangular plots, analogous to the *Melody Triangle*. Additionally the plots of the first and last generations are subtracted from each other to make a ‘correlation heat map’ for each of the orders. These are presented in fig. 5.41

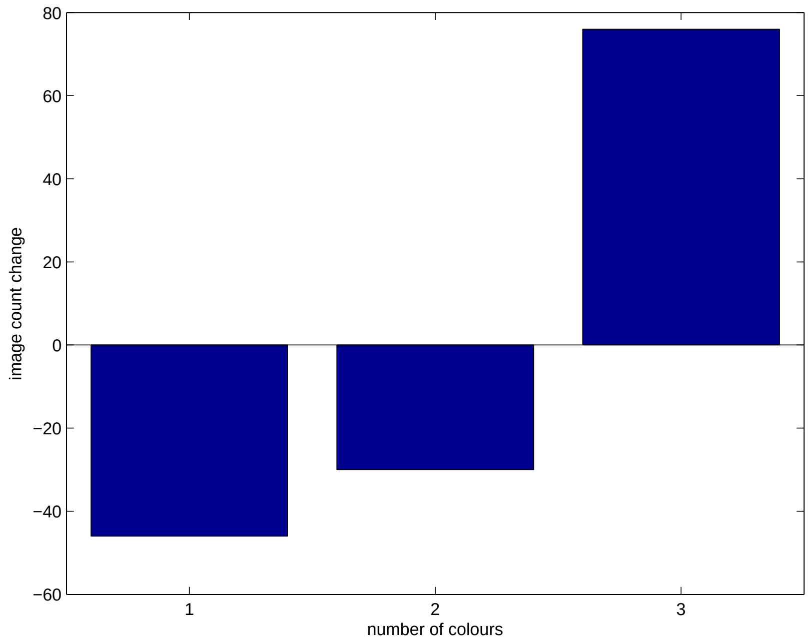


Figure 5.40: Correlation histogram comparing the first and last generation of $EP3$ for the number of colours in the image, $|C|$

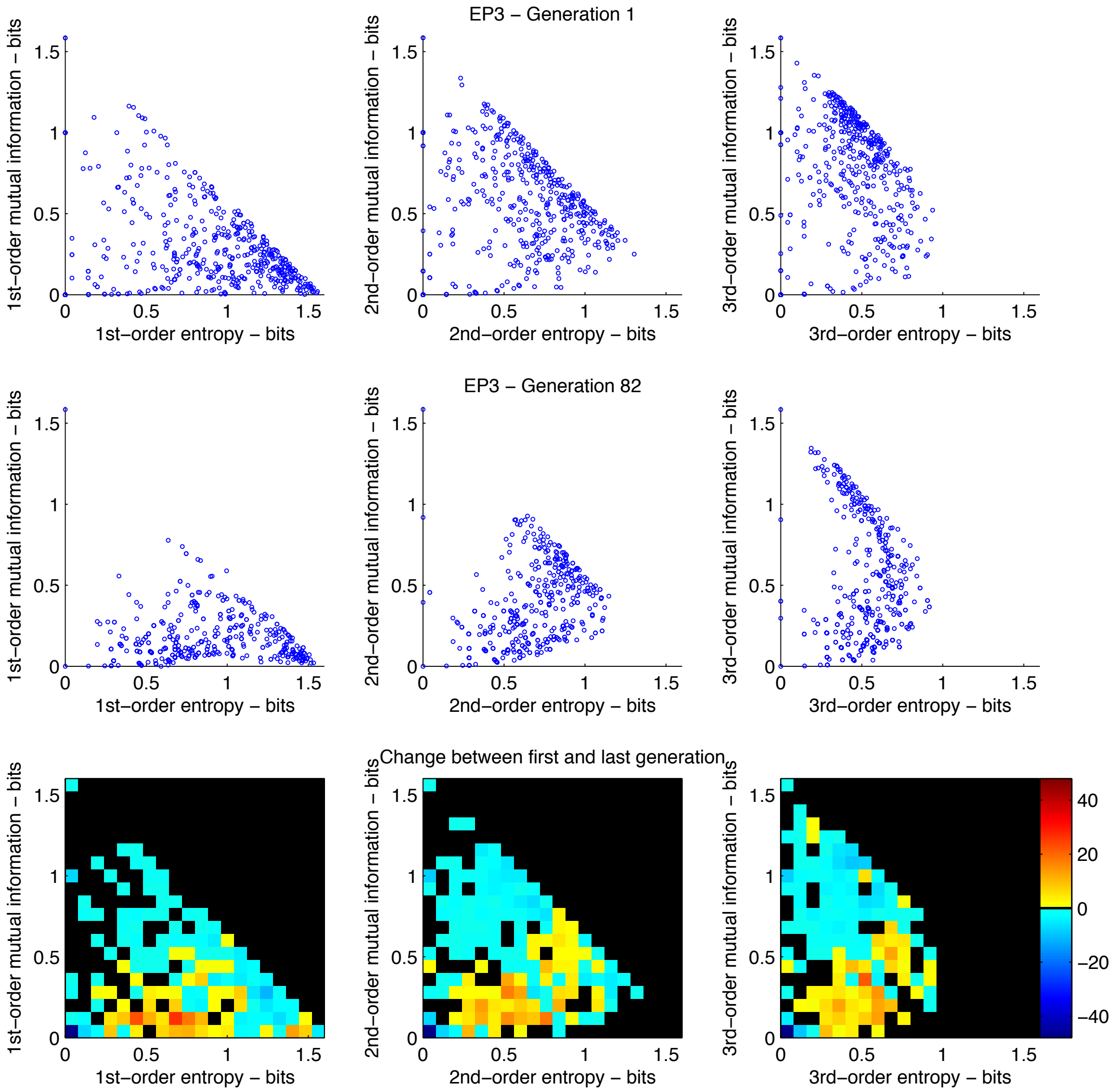


Figure 5.41: Conditional entropy against mutual information for the first and last generations of *EP*3 (row 1 and 2 respectively), with heat maps comparing the two generations (row 3). Column one contains the first-order, $H_{Y|X}$ and $I_{Y;X}$. column 2 the second-order, $H^2_{Y|X}$ and $I^2_{Y;X}$, and column 3 the third-order, $H^3_{Y|X}$ and $I^3_{Y;X}$.

The 1st-order plots represent the most 'local', short term visual properties; patterns only considering the adjacent colours. The 2nd-order and 3rd-order plots consider the patterns of colour grouped in twos and threes respectively, and represent higher level visual properties[16].

[16]As is clear there are significant blank areas in the plots. This is a result of the natural distribution of the Markov chains with regards to the information properties. In the *Melody Triangle*, chains were generated until all areas of the triangle were covered. This could not be done here as the chains distribute differently with regards to higher order measures. Additionally the measures of the *Melody Triangle* were made on a long term theoretical average of the chains and considered a greater number of symbols, than the three colours here, so it was possible to generate a triangle with good coverage. The reason there are no high values for higher order conditional entropies has to do with how colours are grouped together into 'meta-colours' of two or three; there are many more possible 'symbols' (27 different meta-colours in third order), and as the sequences are only 50 colours long, it is impossible to reach a maximum disorder in that

To interpret these heat maps, it is worth considering what the areas of the 'triangle' represent. However some care must be taken with these interpretations, as repeated colours do not appear as multiple discrete events of the same type, but rather appear as one continuous area of colour.

The bottom left, $(0,0)$ corner represents perfect repetition with a single colour, i.e single colour images. They are dark blue in the heat-map, indicating that all such images disappeared.

How far away an image is from this corner represents how many different colours (or groupings of colours for the higher orders) are involved *on average*. At the top left (maximum $I_{Y;X}$, minimum $H_{Y|X}$) are perfect loops of three colours, with the bottom right (minimum $I_{Y;X}$, maximum $H_{Y|X}$) are sequences where all three colours are random, but in equal proportions.

Along the left edge, where conditional entropy is zero, are sequences of perfect repetition. As there are only three colours, in the first-order there are only three possible places to have images with perfect repetition (single image colours, images oscillating between two colours, and images oscillating between three colours, at $I_{Y;X}$= 0, 1 and 1.58 bits respectively). This accounts for the sparseness along the left edge.

Along the bottom edge are completely unpredictable sequences, and the higher the $H_{Y|X}$, the more of evenly the three colours are presented. Areas in between the two edges show sequences that are in between these extremes of predictability.

Two overarching trends are evident. The first trend is that on average images tend to be closer to the bottom left corner at the last generation than they were at the first. Since the ratio of images with three colours increased in the evolution, this indicates that there is a tendency to not have equal amounts of the three colours (or groupings of two or three colours) evenly. The closer to the corner the image is, the more uneven the colour distribution. In the first-oder, this indicates a tendency to repeat the same colours, hence the sparser the image appears.

The second trend is that the less predictable half of the 'triangle' has an increase, with images tending to migrate away from the structured repetitive end of the triangle. However one must be carful to attribute this to a preference for chaotic images, due to the overarching sparseness. Rather this would be more accurately attributed to a *dislike* for highly structured images. Analysing the remaining information measures can help clarify the issue.

H_X is the standard, non-conditional, entropy. It is minimal in single colour images, but maximal if there is an equal amount of all colours. As there is a decrease in this value in $EP3$, despite the increase in the number of colours, this indicates that images with uneven amounts of colours

---

time.

are preferred over images with equal amounts of colour. This is confirmed in fig. 5.42.

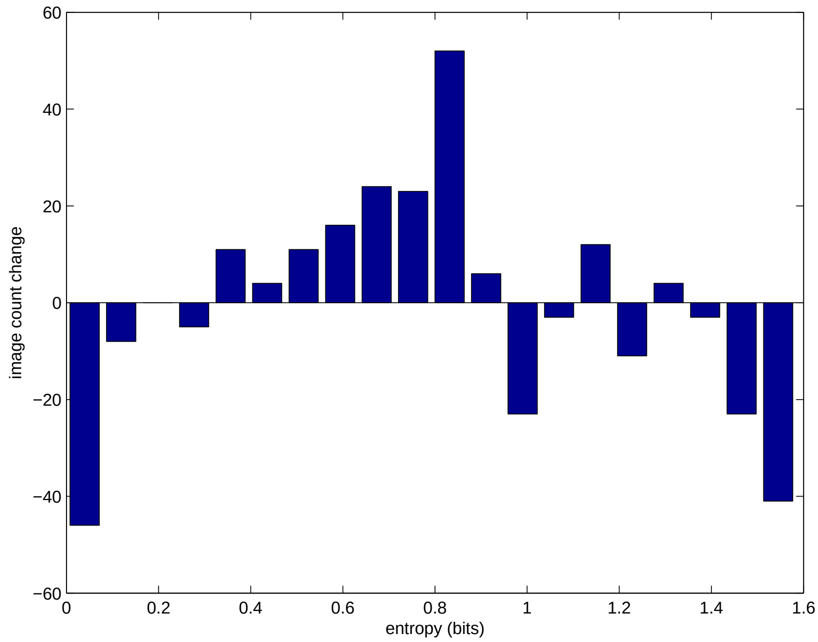


Figure 5.42: Histograms comparing the first and last generations of $EP3$ for the standard (non-conditional) entropy H_X

The following observation can thus be made.

- Observation 23: *Images with unequal amount of colour are preferred.*

Considering the longest unbroken run of single colour, $\max|\vec{c}|$, and the number of colour changes, $|\Delta c|$, the former has significant increase, while the later has very strong decrease. In the first generation, the mean number of colour changes in an image 28, while in the last generation it was down to 18. While the longest runs of single colour went up from 9 to 12. This points to a trend towards greater order, with less busy images with areas of single colour being preferred. A better understanding of these features can be had by considering the correlation histograms in fig. 5.43 and fig. 5.44.

In the longest run of single colour, $\max|\vec{c}|$, the biggest change relates to the decrease in the number of images with short longest runs, that is images that change colour nearly all the time. This would be both images that are very chaotic, or images that have a perfectly repetitive pattern oscillating between colours. It also noted that the single colour images, those with maximally long single colour runs are also disliked. The greatest increase seems to be around single colour runs of about 20, and considering that an image is of 50 colours, these long runs represents sparse

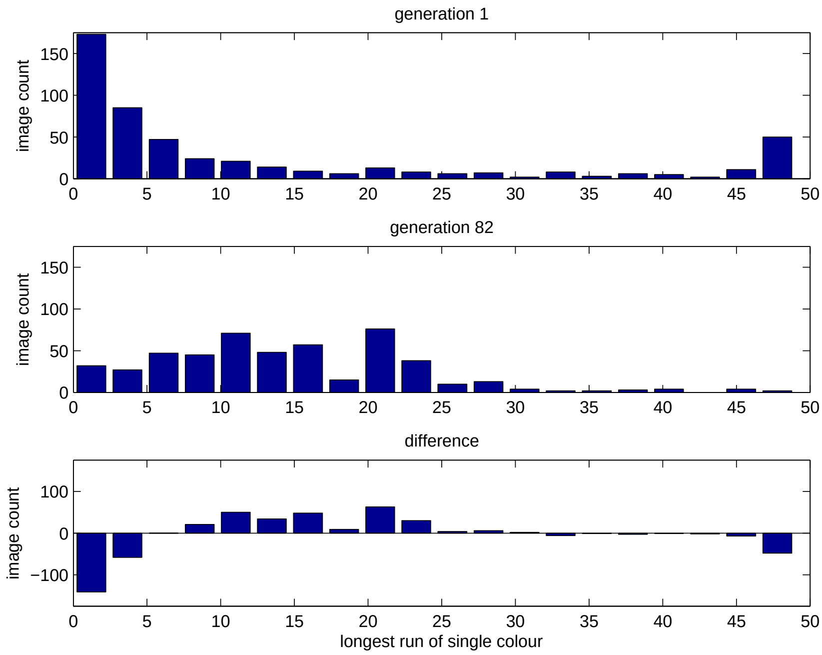


Figure 5.43: Histograms comparing the first and last generations of $EP3$ for the longest unbroken run of single colour, $\max|\vec{c}|$

images.

Similarly considering the number of colour changes, $|\Delta c|$, it also noted that single colour images (with zero changes) are disliked. However images with few changes, peaking at 10 are preferred. Busy images, those with 35 or more are virtually wiped out. The following observations can thus be added:

- Observation 24: *Very busy images are disliked.*
- Observation 25: *Sparse images are preferred.*

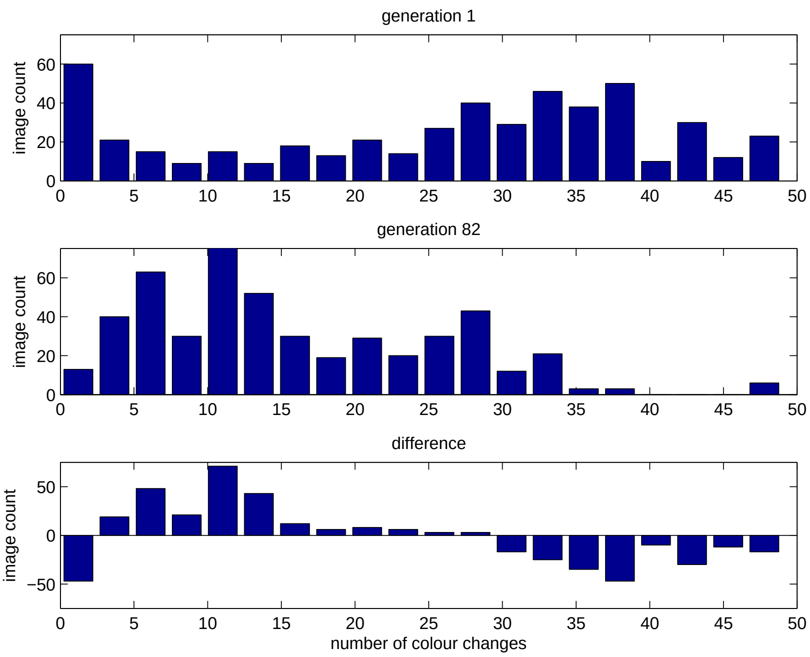


Figure 5.44: Histograms comparing the first and last generations of $EP3$ for the number of colour changes, $|\Delta c|$

5.7 Discussion and Further Work

Table 5.13 provides a full listing of the observations made, and if appropriate any hypothesis they support. The hypotheses were listed previously in 5.3.3.

Table 5.13: List of observations in the analysis of *EvoColour*.

Observation	Noted in	Hypothesis Supported
1 *Dislike for single colour images*	*CP*, *EP*3	
2 *Preference for high lightness contrast*	*CP*, *EP*2, *EP*3	11
3 *Preference for blues; aprox.* h° *275°*	*CP*, *EP*2, *EP*3	1
4 *Dislike for yellows; aprox.* h° *90°*	*CP*, *EP*2, *EP*3	2
5 *Preference for purples; aprox.* h° *300°*	*CP*, *EP*2, *EP*3	
6 *Dislike for green; aprox.* h° *130°*	*CP*, *EP*2, *EP*3	
7 *Dislike for colours of high chroma (highly saturated colours)*	*CP*, *EP*2, *EP*3	
8 *Dislike for perfect repetition*	*CP*, *EP*2, *EP*3	
9 *Dislike for maximum entropy*	*CP*, *EP*2 vs *EP*3	
10 *Preference for images with similar* a^* *values*	*EP*2, *EP*3	
11 *Preference for images with similar* b^* *values*	*EP*2, *EP*3	
12 *Preference for images with similar* C^* *values*	*EP*2, *EP*3	
13 *Preference for images with colours 40° to 90° apart in hue*	*EP*2, *EP*3	
14 *Preference for images with mean low lightness*	*EP*2, *EP*3	6
15 *Preference for* h° *range - 200° to 300°*	*EP*2, *EP*3	
16 *Dislike for* h° *range - 30° to 150°*	*EP*2, *EP*3	4
17 *Preference for high* a^* *values is greater in low entropy images*	*EP*3 vs *EP*2	
18 *Preference for low* C^* *values is greater in low entropy images*	*EP*3 vs *EP*2	
19 *Preference for black and near black is greater in images of low entropy*	*EP*3 vs *EP*2	
20 *Preference for similar hues in images not of high entropy.*	*EP*3 vs *EP*2	8
21 *Some colours on the edge of sRGB gamut are preferred.*	*EP*3	
22 *Three colour images are preferred.*	*EP*3	
23 *Images with unequal amount of colour are preferred.*	*EP*3	
24 *Very busy images are disliked.*	*EP*3	
25 *Sparse images are preferred.*	*EP*3	

As can be seen, all observations made in the control population were also observed in the *EP*3. This suggests that the evolutionary process was able to carry through these aspects of user preference, and as such that any evolutionary biases were outweighed by the volition of the users. This is a form of validation of the evolutionary processes. Had there been observations made in *CP* that were not detected in *EP*3, that would be cause for concern. Conversely that there were observations made in *EP*3 that could not be made in *CP* suggests that using an evolutionary process for the discovery of aesthetic preference does provide advantages over static comparisons, as difficult to detect preferences can be made more prominent in the evolutionary processes.

However each observation in table 5.13 should be not considered a factual finding, and rather be viewed as a question to be asked, and independently tested in future controlled experiment designed specifically for it; the subject of further work.

That some of the key hypotheses suggest by the literature in colour preference were reflected in the evolution is encouraging and expected. However that some the hypotheses presented in section 5.3.3 were not observed, that does mean they are incorrect. The hypotheses derived from quite different experimental situations; some were comparisons of single colours, others compared pairs of colours. As such they are rather different stimuli than the *Markov image*. Further, as mentioned previously a negative correlation of an attribute does not mean it is not preferred, all it means is that some other attributes of the images were stronger. Such situations in *Markov image* comparisons are quite likely, as multiple attributes are tested against each other at the same time.

For example an observation that conflicted prominently is observation 7 vs hypothesis 5 (*Dislike for colours of high chroma (highly saturated colours)*vs *Preference for high chroma colours*). This could mean one of two things: either preferences for high chroma are different in a Markov image than in a single colours, or that there was a bias in the colour mating processes that counted against high chroma values. However as observation 7 was also reflected in *CP*, it is more likely that the preference for high chroma colours does not apply to colours in combination.

Other hypotheses, such as number 9, that suggests there is a preference for images with a figure with 'cool colour' on a 'warm colour' background, and vice versa, is one that simply could not be tested in the evolution; there were not enough images that could be found that were clearly of a figure/ground setup with those colours. The nature of the evolution is such that one does not know ahead of time what hypotheses could or could not be tested.

The observations regarding the preferences of adjacent colours favoured greater similarity, (e.g. observations 10, 11,12 and 20), and if these observations were all simultaneously fulfilled, the colours within an image would homogenise to a single colour (although with different lightness due to observation 2). It is possible that such an image would fare fairly well in comparisons, however almost certainly there are characteristics of aesthetic preference not captured by this set of observations. In particular there is in all likely-hood some optimal minimum distances along the $L^*a^*b^*$components the colours should be apart from each other, but no observations relating to minimum colour distances could be determined. This was difficult to observe in this evolutionary process, as the evolution necessarily manipulates all features concurrently, and the observations made were made on population average trends.

Image fitness could be measured on *conditionals* of the feature values, that is, *given* images with a feature at a particular value or range, how do the preferences for the remaining features pan out?

For example in the observations comparing $EP2$ and $EP3$ (observation 17 to 20) show that there were different colour relationship preferences in high entropy images then in low entropy images. There will be many more such 'conditional preferences'. In particular one could imagine that the proportions of the amounts of colour would determine how far apart it is preferred that these are in colour space. Deng et al.(Deng, Hui, & Hutchinson, 2010) in their study of crowdsourced online shoe designs, found that the relative amounts of colour had an impact on how far apart in colour space the colours would be. In particular, they found that the largest areas of colour would often be of similar hues and close in colour space, but users often chose to have a small element of the shoe design (e.g. the Nike 'swoosh' or the laces) to be of a different contrasting colour. The analysis carried out here was not able to determine if such was the case for Markov images because all the images in a population were analysed together. Further work should be carried out to explore more of these conditional feature preferences. This could be done in the same way that the differences in preferences between high and low entropy images were determined in $EP2$ and $EP3$, by running multiple evolutions with differing setups and comparing their outcomes.

5.7.1 Success

As can be seen, there are numerous observations with respect to colour preferences that were made in the analysis. But should *EvoColour* be considered a 'success'? Was it successful in mapping out the parameter-space of the Markov images with respect to cross-user, objective aesthetic desirability? And could thus the observations of table 5.13 be considered to be attributes of this objective dimension of aesthetic preference? A metric for success was defined in section 1.2.1. It consists of two criteria: $C1$ - *correlation between parameter choices and aesthetic value of the output*, and $C2$ - *sufficient volume of data points and significant trends*.

C1

Were the selections made really with respect to aesthetic preference? When a Markov image is contemplated, the question of wether this image is 'beautiful' is contingent on more than the pixel values on screen. Rather it is the result of the relationship between the subjective viewer and the image, and includes issue of presentation context and the cognitive state of the contemplator.

Further how then a viewer is to articulate that such an pleasurable aesthetic experience occurs is non-trivial; art appreciation is not a process easily measured.

However, the context of contemplation in *EvoColour* shows two images as pairs; the situation is explicitly setup for viewers to indicate a preference. As such the selections are indicators of *relative* aesthetic value: the viewer does not communicate wether or how much they like an image, rather they indicate just wether they like one image more than the other. Such comparisons distill the nuanced and subtle process of aesthetic contemplation to a binary selection. How is one to determine if this is a valid thing to do?

A core assumption here is that there is some sort of correlation between preference judgments, and a notion of what is of aesthetic value. It is impossible to peer into the brains of the participants and determine if this really is the case, however the data does point to patterns in these selections. There is evidence of a cross-user *consensus*, of a global agreement over which images are to be preferred over others. This is evident even just by looking in the control population, *CP*, where it was clear that some images are consistently selected over others.

That there exists such a consensus across participants suggests that there is an objective element to user preferences. We see this kind of consensus of aesthetic judgments in the 'real world', where aesthetic responses within groups tend to correlate, and there is a recognition that some aesthetic judgments are more apt than others(McMahon, 2011).

This would seem to suggest that there is a correlation between the binary selections of preference and aesthetic value. This fulfils part of the requirement for criteria $C1$.

It must kept in mind however, that *EvoColour*, like many crowdsourcing works, suffers from issues of medium-specificity. As the participants are all out 'in the wild', there is no way to control for differences in the viewing environments and contexts in which they interact with the system. Most colour perception research is carried out in very controlled conditions with carefully calibrated monitors. With crowdsourcing this is not an option.

In *EvoColour*, the display used can vary wildly, as any device that can render a web-page can be used to participate. As such, viewers will be using screens of different sizes and capabilities, as well as participating in wildly differing environments and context. For instance, a viewer interacting with *EvoColour* on a small mobile screen while on a crowded bus will have a very different experience to one interacting with it in on a large high-definition display while alone in a dark space. Every computer display is different and renders colour differently. Further, every setting of a computer screen, be it brightness, contrast or saturation, will all also change how the colours are rendered.

There is also no way to control for the environment in which these preference judgments are made. Some viewers will be looking at the monitors in rooms filled with ambient daylight, others will be making their preference judgments at night. Environmental conditions could affect how these Markov images are perceived, and consequently will have an impact on the data gathered. Environmental conditions in other modalities could also conceivably affect results; could judgements be affected by the sonic properties of the environment? Will the data collected differ if there is noise, music or silence in the space? Will the results be different if the room is warm or cold? What about if the space is crowded or empty?

All these factors will add noise to the collected data, however the difficulties do not end here. In addition to environmental and display variabilities, there will be noise introduced by variability in the cognitive state of the participants. Levels of fatigue and mood could conceivably affect the comparisons; a participant may make different judgements when they are wide awake and in jovial mood, from when they are exhausted and in a foul mood.

Further, the process of participating in *EvoColour* over time will likely have an effect on the judgments. A participant is not likely to be making preference judgments the same way during their first comparison, as they are after having done a thousand comparisons. Also preceding context could influence the collected data; if a participant has seen a dozen blue images, and then suddenly sees a red one, it is conceivable that its novelty may affect how it is judged.

What then to make of all these issues? Every issue mentioned above is an uncontrolled variable that will add noise to the collected data. However there is one even greater source of noise: the subjective difference between individuals' tastes. If aesthetic judgments have both an objective and a subjective element, then just as the medium-specific context for each participant adds noise to the system, so does the variability and idiosyncrasies of individual tastes.

But this is precisely one of the advantages of evolutionary processes: they are able to handle noisy data. In the same way that the evolution pulls towards an averageness of fitness, and this averageness removes idiosyncrasies of taste and only represent the objective aspects of the judgments, the averageness of the evolutionary process will also be removing the noise introduced by the variabilities in medium and context of display.

Every individual has their own set of personal preferences when engaged in *EvoColour*, their own 'fitness landscape' of sorts, and how this landscape differs from that of others represents the subjective elements of their taste. However, this landscape is dynamic and shifting all the time; it will be affected by the cognitive state of the viewer, it will morph as they gain experience, and it will shift in line with medium-specificity issues mentioned above.

The evolutionary process is akin to a low-pass filter, where the high frequencies and variabilities are filtered out, and all that is left are what is common to most viewers: the objective elements of the aesthetic judgments, with the medium-specificities averaged out.

The next question then becomes, did the evolutionary processes of *EvoColour* drive the parameters for populations of Markov images in way to correlated with the aesthetic preferences of the participants? In other words, did the evolutions really make the populations 'fitter'? This was verified empirically with the fitness comparison selections that the users were made to do every four selections, comparing images from the latest generation with that of the first. As was discussed previously in section 5.6.2, the mean fitness of the populations increased across the generations.

All told, it is the present author's opinion that $C1$ has been met: there is a correlation between the binary judgments of preference and the aesthetic value of the artefacts, and the evolutionary process was able to carry through the mean force of aesthetic preferences, as each generation became 'fitter' than the last.

C2

Turning now to success criteria $C2$ - *sufficient volume of data points and significant trends*. There were 45499 selections done as of May 23rd 2015. Could this be considered enough data points relative to the parameter space of the Markov image? The analysis of *EvoColour* succeeded in making a large number of observations, significantly more than was initially anticipated. There were initial concern that the Markov image would have either too great a potential complexity for the evolution to lead to identifiable trends with respect to the number of selections that could be gathered, or that it would be too difficult to learn anything new in analysis. It seems the balance was about right; the state space of the Markov image is suitable for evolutionary processes, and lends itself well to analysis.

The ability to draw so many observations suggest there were enough data points, fulfilling criteria $C2$. As it would appear that both criteria $C1$ and $C2$ were met, it seems safe to suggest that *EvoColour* was a 'success', and the observations in table 5.13 could be said to be attributes of objective, cross-user preferences for these images.

However, the plots of the parameter values over the generations did not indicate that 'final' values were arrived at in *EP*3; in all likely hood, the evolutions could progress for many more generations before fitness would cease to improve and the attributes of the images stabilised. Even though it could be said the images in the final generation in *EP*3 look on the whole more like each other in the last generation than in the first, there are still many varieties and differences in the images.

Had such a plateau been reached, it would not mean that the attributes of an 'archetypal image', fitter than all the rest, had been found[17]. Rather it would point to a limit in the mating algorithm, where mating fit parents would no longer create children fitter than the parents. Further practical considerations (in particular the issues of medium-specificity discussed above) that introduce noise to the preference selections, could come into play. However these limitations did not seem to have yet been reached in the evolution.

Although a considerable number of selections were made, it would have been even more insightful to get many more; to have the evolution approach these limits of performance. Improving on the gamification elements, such as providing a 'current score' or 'ranking' on screen could could help achieve this.

5.7.2 Further Work

Families of Interest vs. The Global Average

It did not appear that the evolutions in *EvoColour* had reached a peak in the fitness landscape; the mean attributes of the images in the populations were still evolving. However, as the images increased in fitness, the populations of images became increasingly homogenised. This is evident for instance in observing that the populations as a whole were tending towards an overall blueness. It was these average trends that made it possible to extract the observations listed in table 5.13. It is expected that should the evolutions continue for much longer, then the increase in similarity of the images would continue. As is the natural tendency of evolutionary algorithms, a convergence towards a peak of fitness would occur.

However it is not known if this peak that the populations approach is the highest peak in the

[17] That such an image could be said to even exist is contentious. And even if on the whole blue is the 'world's favourite colour', this does not mean that the (hypothetical) fittest image is necessarily even blue.

fitness landscape. Moreover, the highest peak – if such a peak can truly be said to exist – is not necessarily the only point of interest. All generative processes posses families of interest in parameter values, and these may lie in distinct and separate areas of parameter-space. It is known, for instance by looking at the most popular images in the non-evolving control population, *CP*, that there are popular images that are not blue. It is of value to be able to seek out these other sets of parameters values; those other smaller peaks of fitness.

The first way to seek out these other peaks of interest, would be simply to re-run the evolutions. If the experiment was re-launched, seeded with a new population of random images, it is not known if the evolution would necessarily go towards the same fitness peak as in the previous run. However having observed the strong tendency to push towards the blue area of colour-space in all the evolutions, it is likely that something similar – with respect to overall colour certainly – would happen again.

What if one is out to find the parameters that would make the best red images? Re-running the evolutions, no matter how many times, is not likely to happen to converge towards a population of red images. As such the evolutionary algorithm that drives *EvoColour* in its current form may not be well suited to finding many disparate peaks. However there are two ways that *EvoColour* could be modified to explore an area of state-space specifically.

The first is by curating the seed population of images. For instance, if one is out to find the parameters that would yield the most popular Markov images that are mostly red, the population in the first generation could be seeded with exclusively red images. This would ensure that the evolution is focused in this area of colour space, and the evolution may then converge on a local maxima of red images[18].

An alternative would be to prevent some attribute of the images from being affected by the evolutionary process. For instance, the mating and mutation algorithms could be modified to not affect the b^*component of the colours of the Markov image. This would disable the tendency of populations of converging towards the blue areas of colour space, and thereby allowing other attributes of these images to come to the for, revealing differing areas of the fitness landscape.

Mitigating Biases

There is no guarantee that the evolutionary and mating algorithms of *EvoColour* were optimally designed; almost certainly there are inherent biases in the mating processes. In the same way

[18]One may also need to ensure that mutations do not introduce new colours to the pool

that $EP1$ and $EP2$ had a bias for images of high entropy, there will be other, undetected biases in $EP3$. Although considerable care and effort went into designing the evolutionary algorithms of *EvoColour*, their designs embody a set of beliefs and assumptions. For instance, they contain the assumption that global average colour and the relative geometry of the colours in colour space are related to preference. These were informed assumptions, as the literature on colour preference and harmony surveyed in sections 5.3.1 and 5.3.2 suggest. However there will have been many ways to interpret and embody this knowledge into the algorithms; alternate designs or parameter selections for the mating algorithms may well have been able to improve fitness faster.

For example, the colour mating algorithm of *EvoColour* translated the relative geometry of the images' colours with respect to their $L^*a^*$and b^*components. However given that preference is often related to hue(Eysenck, 1941; Schloss & Palmer, 2011; Ou & Luo, 2006), even faster fitness improvement may have been achieved if the mating algorithms were designed with respect to the colours' $h^\circ$and C^*components.

Similarly there could be improvements with regards to the mating of the images' Markov chains. As the mating algorithm of $EP3$ simply takes the Markov chains of one parent or the other, the evolving populations' chains are highly dependent on those that exist in the seed populations. There is thus an inherent bias with respect to how the seed population's matrices were initially created. As was shown in the analysis of the information properties of the evolved populations, shown in figure 5.41, the starting generations had a distribution of transition matrices not evenly spaced out in information space. A possible improvement would be to seed the starting population with Markov chains evenly distributed in information space.

Because the evolving populations increased in fitness as the evolutions progressed for each generation (as discussed in 5.6.2), there was no indication that these biases engendered a ceiling of fitness. However it is likely that they will have throttled the rate of fitness increase.

The approach taken for designing the mating algorithms for *EvoColour* described in section 5.4.3, includes a process that tested likely features of interest against simulated evolutions. These simulated evolutions checked that for any given feature, that it was possible to evolve images towards both maximising and minimising these features in a population, as well as checking that the features would not homogenise or drift radically under a simulation doing random selections. This is an optimisation problem that was done manually here, however it is conceivable that such a process could be automated.

This process could be further improved if additional, more informative features were employed to guide the design of the mating algorithm.

Additional Analysis Features

The analysis and mating algorithms would have been even more fruitful had some more features been identified. In particular, some omissions that in hindsight would have been desirable relate to capturing the proportions of the various colours. The overall (non-conditional) entropy hx does this, but more detail would have been useful. In addition to the 'longest unbroken run of single colour', $\max|\vec{c}|$, capturing the 'average run of single colour' and the 'minimum run of single colour', would have enable a more nuanced analysis with regards to colour proportions and arrangement.

Additionally it should be possible to see how well images conform to *Zipfs's law*[19], which has been advocated by number of researchers to relate to aesthetic pleasure (as discussed in 2.6.3), by measuring the amounts of each colour rendered and calculating their relative distributions.

That both images of maximum disorder (observation 9) and of perfect repetition (observation 8) did not fare well in the evolution fits well with the hypothesis of preference for *effective complexity* suggested by Galanter (Galanter, 2003), as well as the views of many other theorists discussed in section 2.6.2. However the observation only considers the extremes, and does not say anything about how much order or disorder is preferred, nor is it know how exactly this disorder should be measured.

The information theoretic measures used, such as the conditional entropies and mutual information, did provide some insights, but the difficulty is that in these measures repeated colours are considered discrete 'events' of the same type, whereas they are not *perceived* as discrete events in a Markov image, but as one singular event. Some researchers have been known to use *jpg* or *png* compression file sizes to estimate the complexity of an image (e.g.(Rigau et al., 2007; H. Yu & Winkler, 2013)). It would be interesting to see what observations of such a feature could have added. With a measure of complexity, additional proposed aesthetic measures could be tested, such as those inspired from Birkhoff's measure(Birkhoff, 1933).

A measure of 'directionality' could be developed, by measuring the difference in the information measures depending if the image was 'read' from the inside out (as they were generated here) or from the outside in. If a significant difference in the measures correlate to preference,

[19]Where distributions that conform to a $1/f$ statistical power law are suggested to be more pleasurable

this could yield insight into how viewers process an image spatially.

Practical Applications

An intriguing area of further work is to apply the knowledge gained from this study to design and in the construction of tools. The observations in table 5.13 are in and of themselves valuable, as they can form the basis for design heuristics. These could be manually considered by designers, and could be used as design guides where ever colour combinations are to be applied. As mentioned in section 5.3.2, there already are numerous design guidelines for colour 'harmony', however not all have empirical evidence to support them.

An exciting potential application of this research is to incorporate these heuristics into the construction of smart design tools. Such tools could provide an objective aesthetic evaluation of an artefact under construction, providing suggestions for ways to improve the colour aspects of the designs. It is even conceivable these heuristics could inform the construction of software that could stand in and autonomously make creative decisions that are likely to appeal to many.

The Archetypal Image

The original artistic motivation for the creation of *EvoColour* was simply to create images of beauty. But more specifically, and motivation was to try to find the image (albeit within a restrictive set of constraints) that would appeal to the most people. What would this 'archetypal image' look like? Can it really be found? If it does exist, will it be bland and un-offensive to appeal just enough to many different people? Or will it have some striking features and resonate strongly with all viewers?

It would be intriguing to use the observations to defined a *computational fitness function*, and then run an automated evolution to evolve the 'archetypal image' that best matches those goals. However it is unknown how to weigh the importance of each of the components of such a fitness function. It is unlikely that each observation is as relevant to overall fitness as each other, and a different weighting would yield a different result. Often in multi objective optimisation, *Pareto ranking* is employed to rank individuals where the objectives conflict with each other(Galanter, 2012), such that it is impossible to improve the score of one objective without decreasing the score of another. Interestingly, there does not seem to be many conflicting observations in table 5.13[20], perhaps indicating the evolution was on course towards some form of global homogeni-

[20]Perhaps with the exception of observations 3 and 5, however an image could be in between purple and blue (e.g. at $h°287$).

sation.

Another approach would be to use mean, or perhaps the mode[21], of the feature values to define an image. Such an image would represent the most typical average image in the population, however it is unclear if *averageness* would yield the highest fitness[22].

To use the results of the evolutions of *EvoColour* to synthesis an 'archetypal image', and evaluate its fitness is further work. However a very corse estimate is provided. On the *EvoColour* home page, as shown in fig. 5.1, is a display of the 'most popular image'. This image was determined simply; it is the the image in the latest generation that has won the most comparisons. This is not an archetypal image, but is certainly a very fit image, and conforms with most of the observations in 5.13. This 'most popular image' as of May 27th 2015 is presented below in fig. 5.45.

The image reflects the closest that could be found to an 'archetypal image', but who's image is it? Like the other works in this thesis, there is an ambiguity in authorship and attribution. The decisions to use circles, to use Markov processes, to use evolutionary processes with crowdsourcing all were done by the present author, in part motivated by a scientific curiosity, in part motivated by an artistic drive. Yet the parameters that went into this Markov image were driven by the volition of participants, and this image simply would not exist if it were not for the numerous participants and thousands of preference judgments made. There is no simple answer here, but neither an attribution to the present author, nor to the anonymous 'crowd' seem correct. Perhaps with this kind of crowdsourced artefact, attributions are best left unspecified.

If the evolutionary process does successfully remove all subjectivity, and the relationship between an image and its fitness correlates solely with respect to the cognitive processes that are common to us all, it is possible that there is stable objective fitness landscape. This would imply that there exists an image that is at a higher peak than the rest, an 'archetypal image'. However, the task of finding it is akin to a difficult optimisation problem. Within the highly dimensional space, there is clearly not just one peak, and knowing wether one lies at a local maxima is impossible. The analysis of the features and the observations made in this research are akin to explorations in this landscape as we look for higher ground. The lists of observations describe how things tend to be on this higher ground, but there is a fog above, and there we do

[21] if the real number features were quantised into groups, such as in the histograms.

[22] 'Averageness' has been noted as being an attractive quality in human faces, however Perrett et al. in their studies of the attractiveness of human faces, found that although the 'average faces' are considered attractive, they were not the most attractive(Perrett et al., 1998).

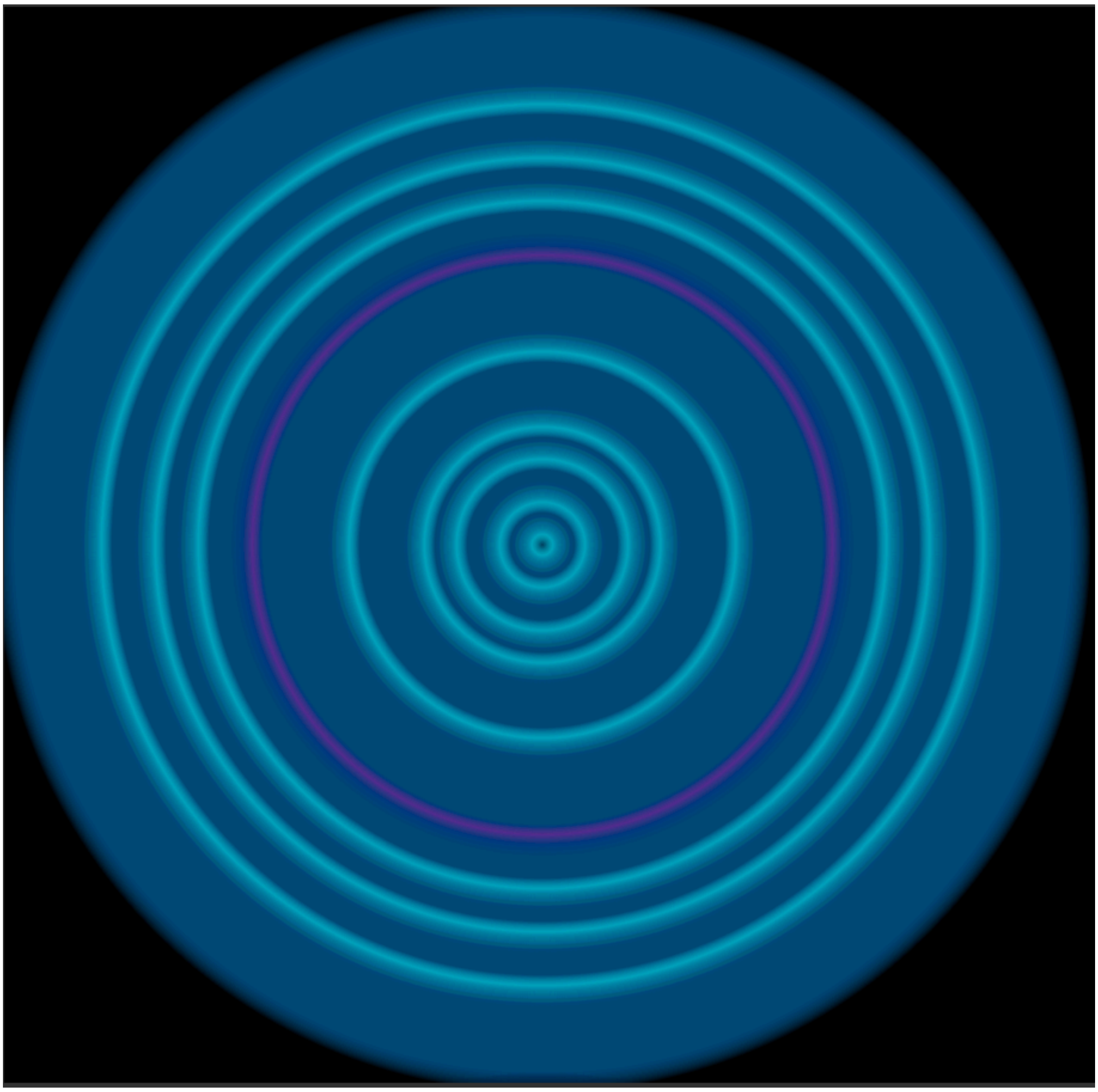


Figure 5.45: Image in generation 82 of $EP3$ which had won the most comparisons throughout its lifetime, as of May 27th 2015

not know what lies, or where the highest peak may be.

This chapter presented *EvoColour*, a crowd-sourced interactive evolutionary system for the evolution of patterned colour images, called *Markov images*. It was contextualised in previous literature on colour models and colour preference research. The design of the evolutionary algorithms were detailed, including a novel colour palette mating algorithm that sought to maintain the geometrical relationships between the colours in colour space.

An extensive quantitative analysis of both evolving and a non-evolving population was carried out. This resulted in numerous observations on aesthetic preferences for colours and colour combinations. Some of these supported previous work in colour research, some were new observations. Additionally observations could be made on the preferences for arrangements and relative amounts of theses colours. It was shown how many more observations could be drawn from the evolving populations than just the non-evolving population, suggesting that the approach taken here – using interactive evolution on generative systems with state-spaces of modest size – is a fruitful approach for empirical aesthetics. Finally a discussion of the findings was provided, and avenues for further work outlined.

In the next, and final chapter, the three studies of this research – the *MelodyTriangle*, *Keyebernates* and *EvoColour* – are discussed in a broader theoretical context.

Chapter 6

Conclusion

A key concern of this research is to elucidate the nature of algorithmic creation in the arts. In algorithmic creation, an artist, designer or composer uses some form of abstract formal procedure – a generative process – to assist in crafting forms and patterns. These forms can manifest themselves in any media; from patterns on a quilt, sculptures, printed images, musical scores, digital artworks or even architectural constructions and robotic devices. Evidence of the application of algorithmic thought are found across cultures, and across the centuries; generative processes are ubiquitous in both music and the arts, and some theorist even suggest that it is 'as old as art itself'(Galanter, 2003).

Algorithmic thought affords new creative possibilities. When an artist or composer applies formal procedures to a creative act, the space of possibilities is defined by the logic of the procedures. This logic can make some previously difficult to attain forms come to be within reach of the artist. There is a subtle interrelationship, an *equivalence* even, between style and generative process; they are kinds of heuristics that reduce the state of possibilities, and define *constraints* within which a class of works can be generated(see 2.1).

Some formal procedures can be explicit, with the instructions spelled out, sometimes in the form of computer code. But they can also be spelled out in other places; for instance such procedural instructions can be found in religious texts defining ritual practices or architectural construction(Trivedi, 1989; Datta, 2010; Eglash, 1999). Alternatively these formal structures may be implicitly encoded in collective cultural knowledge, and are manifested as styles, norms and genres.

A generative process can define a set of *parameters* that allow for the manipulation and control of the generated form. This happens when the generative process is *under-specified*, and the remaining elements of the system are defined in a separate step to the definition of the core algorithm. This thesis focuses on the relationship between parameters, their impact on form and ultimately on the aesthetic value of the perceived forms. Generative designs come with difficulties and peculiarities; the relationship between a parameter and its intended result is sometimes hard to predict, and when parameterised also hard to control, as there can be numerous layers of causality between the designer's action and its consequence. One of the aims of this research is to explore approaches to facilitating this interaction.

Parameters, due to their numerical nature, can be easily manipulated and controlled, and as such lend themselves well to scientific analysis. This research employes generative systems as a framework through which questions of aesthetics value, taste and subjectivity are raised and explored. To main approach is to, through practical studies, elicit parameter selections from individuals with respect to judgments on the outputs of a generative process. As choosing parameters for an artistic or musical generative system involves making a judgment of value, the analysis of the parameter selections can yield insight into the nature of such aesthetic judgments. These insights can help elucidate the relationship between objective and subjective dimensions of aesthetics, and can also yield practical applications with regards to the design of creative systems for the automatic generation of aesthetic objects or their evaluation.

The core aims of the thesis are thus reiterated:

- To elucidate the nature of algorithmic creation, and the relationship between generative processes, parameter search, and aesthetic value.
- To explore novel mechanisms and interfaces for parameter discovery and the exploration of the aesthetic possibilities of generative processes.
- To discover cross-user trends in aesthetic preferences in visual patterns and musical preferences.

Three practical studies were carried out in this research, each embodying differing approaches to the exploration of the aesthetic potential of generative processes. The first practical study was the *Melody Triangle*, a musical interface that provides a mapping between the parameters of a stochastic generative process, the Markov chain, and the phenomenal predictability of its output.

It enables the user to navigate the musical possibilities of the process with respect to high-level measures of subjective predictability. As a mobile phone application, the favoured parameters for these stochastic processes were crowdsourced.

The second practical study, *Keyebernates*, is both an art-work, as well as a novel eye-tracking based framework for the exploration of the possibilities afforded by generative designs. The interaction between the viewers' gaze patterns and the system engender a fluid navigation of the parameter space of the visual forms. This is the result of a delicate balancing of randomising noise processes against the volition of the viewer, their gaze a guiding and organising force.

The third practical study, *EvoColour*, is an interactive evolutionary system that leverages the 'wisdom of the crowds' to find the parameters for a generative system that renders abstract images of concentric circles. It collected tens of thousands of aesthetic judgments from hundreds of users world wide, interpreting these to find pleasing colour combinations and arrangements of these colours, in a never-ending evolution towards ever more compelling images.

The three systems designed as part of this research are frameworks that 'wrap around' fundamentally simple generative processes at their core. These processes – the interlocking stochastic melodies of the Markov chains in the *Melody Triangle*, the interweaving circles of *Keyebernates*, and the patterned colours of *EvoColour* – although relatively simple, still pose challenges to parameter discovery. In particular, they exhibit a non-linear relationship between parametric distance and phenomenal similarity – due to the unpredictability of emergent form – and exhibit a large gamut of possible outputs. All three systems created as part of this research provide an alternative to manual parameter search for these processes; they are systems for *parameter search and discovery*.

Each study collected data from the participants in the form of parameter selections, and the data was then analysed to look for patterns in the selections. A metric for 'success' was defined (see 1.2.1) for determining if these extracted patterns could yield valuable insight on the aesthetic preferences of the participants. The success metric consisted of two *criteria*:

C1: *correlation between parameter choices and aesthetic value of the output*

C2: *sufficient volume of data points and significant trends*

Only *EvoColour* was deemed 'successful' according to this metric; it found numerous patterns in the preferences judgments of its users with respect to colour and colour combinations in its core generative process. Although *Keyebernates* and the *Melody Triangle* did not meet the criteria, they possess value as novel creative interfaces in their own right.

The three studies will be returned to and placed in wider theoretical context later in this chapter. First the question of the nature of generative systems and algorithmic creation, and their place in the history of art and music (the subject of chapter 2), is re-visited.

6.1 Generative Process in Music and Art

This thesis presented historical review of generative processes in art and music, identifying common themes and theoretical concepts. It began by tracing the history of algorithmic thought in Western music (section 2.2.1), from the beginning of notation straight through to modern day laptop live-coding performances. The narrative ark of formal processes in music builds on the inherent *codifiability* of music, as manifested in the quantisation of sound in time, through the *lattice sonics*(Wishart & Emmerson, 1996, p. 8) of the musical score. There is an equivalence between the structures and restrictions of musical style, and those of generative processes; musical traditions help guide the composers output, just as procedural formalisms can. Wether a work is to be called 'algorithmic' or generative, is really a question of how visible and externalised the formal procedures are, and indeed of wether the composer declares their use of formal procedures. That is, algorithmic creation can be *explicit*, such as when the formal rules are specified, or the formal processes are *implicitly* present in the creative process.

The history of Western music, and the evolution of styles, can be seen as an increase in the complexity of the procedures and constraints on the 'lattice sonics' of music. When Schoenberg conceived serialism, it was a recognition that tradition was indeed such a set of formal structures, and that they could thus be replaced by another set of formal rules. Since then many composers, such as the so called 'minimalists', would have a more explicit and transparent relationship to generative processes in music composition(Reich, 2004, p. 35).

In the visual arts (see section 2.2.3), pattern creation is evident across cultures. The *definition* of a pattern can be understood as an algorithm for a generative process. The motivations for the use of generative processes are wide and varied. Some are driven by spiritual concerns, this is the case for generative Islamic tessellations(Marks, 2006) and the generative architecture of some Hindu temples(Trivedi, 1989), as well as the work of artists such as Agnes Martin(af Klint et al., 2005, p. 24). In other cases the uses of such processes were simply pragmatic, as is the case for industrial manufacturing. In yet other cases, artistic exploration considered the generative process itself as the subject of the art, as exemplified by the work of Sol Lewitt(LeWitt, 1967).

A key recurring theme of generative art and music is that of *indeterminacy*. Indeterminacy was explored by avant-guard composers such as John Cage, as well as studied by scientists such as Shannon and Weaver in their research into information theory. Information theory would prove highly influential to both artists and composers, particularly as computers became more widely used. Stochastic procedures were explored by the first pioneers of both computer art and music.

Computer music (see section 2.2.2) would have broadly two concerns. Some scientists used computers to carry out 'musicological' explorations, whereby the existing structures of music were encoded to generate compositions in particular styles (such as in Hillier's *Illiac Suite*(Hiller & Isaacson, 1959)), whereas other composers, such as Xenakis, saw information theory and computation as an opportunity to create new, never before heard sound worlds(Xenakis, 1992). Accordingly, generative process would not only be used to create musical scores, but with the advent of sound synthesis, be used to generate new mechanisms of sound production.

The earliest computer art (see section 2.2.4) would also draw heavily on stochastic processes and information theory, the works of the earliest pioneers contextualising their work in an 'information aesthetics'; an approach to aesthetics that is entirely objective and scientific(Nake, 2005). In addition to information theory, the field of cybernetics would prove to be highly influential to generative artists. It brought with it new understandings of the complex, dynamic systems that underly natural processes, and in tandem, techniques to simulate and study these process would be appropriated by computer artists and composers. These techniques include neural networks, L-systems, cellular automata, agent-based systems, and genetic algorithms, the later inspiring a whole sub-branch of artificial intelligence called 'evolutionary computation'. As these techniques matured, they would be grouped together into a field called *artificial life*, or *a-life*, a branch of both scientific and artistic enquiry concerned with the simulation and synthesis of living things. A-life art represent an extreme, and kind of culmination of generative art. Here artists explicitly seek to make works that 'make themselves', it is an act of 'creating creation', or *metacreation*(Whitelaw, 2004).

Perhaps the appeal of generative processes come from their tendency to appear to provide a sense of self-organisation. Nowhere are generative processes more prominent than in nature; life itself can be viewed as a kind of emergent self-organisation. Perhaps it is this that drives artists and composers to generative processes, the *emergence* they engender serving as a reflection on the processes of nature (see 2.5).

6.2 Unpacking and Contemplating Generative Processes

But what is a generative process, really? This question is the topic of section 2.3, and is approached through formal definitions and concrete examples.

A generative process can be viewed as a set of instructions on low level elements that engender an *emergence* of higher level form. However there are always multiple interpretations to any given situation; a generated artefact can be described at different levels of abstraction, and so can the instructions the constitute the generative process. Further what even constitutes a 'system', and where it ends and its 'environment' begins, is not always clear cut and is ultimately a subjective consideration.

Some generative processes can be 'under-specified', leaving aspects of the process to be determined in a separate step. This is parameterisation (see section 2.3.2); *parameters* are input into a generative system to complete its definition, and then is an output generated. The generative process can be seen as a set of constraints on the possibilities of a system, and how any the output of the system is constrained, is defined by the range of possibilities of its parameters. However just as a generative process constraines the range of possible designs, it also makes these designs *discoverable*; the generative process *affords* the discovery of certain forms. Constraints and affordance have an equivalence.

Like parameterisation, non-determinism is another way to 'under-specify' a generative process (see 2.3.3). By applying stochasticity to a generative process, such as Markov chains, to a process, designers relieve themselves of some of the decisions to be made. Stochastic processes can themselves be parameterised, and techniques of information theory can be used to measure degrees of uncertainty in a generative process.

The causal link between input parameters and the generated output can be non-linear and difficult to predict, even with simple generative processes (see 2.4.1). This is due to the emergence that occurs as low-level elements interact or intersect, bringing forth higher-level patterns and forms. Therein lies some of the difficulties faced by algorithmic composers and designers; there is little that can be gleaned with regards to the subjective phenomenal properties of a generated output by observing the representation of the state a system (the state-space), or by observing its input parameters (the parameter-space). Parameter values that are close to each other in parameter-space can yield phenomenally different results, how then is a designer to approach the search for effective parameter values for a creative or artistic context? This is one of the key questions that this thesis seeks to address.

This problem could be approached by defining a set of *features* and use them to describe the generated artefact (see 2.4.1). Features can provide objective quantitative measurements of properties of an object. However knowing which features are relevant to any particular task at hand is non-trivial, and particularly so if one is to attempt to arrive at some measure for the aesthetic value of the generated artefact. Many researchers have sought to define features that act as measures of aesthetic value, however such measures, such as Birkhoff's *Aesthetic Measure*(Birkhoff, 1933) or *Zipf's law*(Newman, 2005) do not usually stand up to scrutiny(McWhinnie, 1968; R. Davis, 1936).

Before one can talk of aesthetic value, an understanding of the perceptual mechanisms at play when humans contemplate a generated artefact need to be considered. A simpler problem than measuring aesthetic value, is simply to seek to arrive at a way of measuring *similarity* between generated artefacts(see 2.4.1). If one cannot even measure how perceptually similar two artefacts are to each other, how then one is to even begin to determine if they posses valuable aesthetic properties?

6.2.1 Conceptual Spaces

Peter Gärdenfors argues "judgments of similarity reveal the dimensions of our perceptions and their structures"(Gärdenfors, 2004a). However neither considering differences in the variables of a system's state, nor its parameters or features values necessarily leads to a way of reliably measuring similarity between artefacts. Gärdenfors' theory of *conceptual spaces*(Gärdenfors, 2004b), a framework for representing concepts based on geometrical structures(see 2.4.2), is an approach to the problem.

Within conceptual spaces, similarity is modelled naturally by relating proximity in the geometrical structures to similarity. A number of *quality dimension* together define a *conceptual space*. Each quality dimension represents the various 'qualities' of an object. The dimensions form the means by which properties are assigned to objects, and serve as a framework for identifying relations between them. A quality dimension can correspond to how sensory receptors receive information from the world, such as brightness, temperature or pitch. However they can also relate to abstract properties, such as notions of order and disorder. Quality dimensions can be learnt (such as when children begin to distinguish between the height and volume of an object), and additionally can be ‘discoveries’ of science, such as Newton’s distinction between mass and weight, or his discovery of the colour circle that defines the quality dimension of hue as being circular.

With the quality dimensions of a conceptual space in place, it becomes possible to relate the similarity of objects with respect to their proximity in the conceptual space; the closer they are, the more similar they are. Conversely, the further apart they are, the more dissimilar they are. A convex area of the conceptual space can be a ‘natural concept’, and Gärdenfors proposes that natural concepts often coincide with linguistic ascriptions. For instance ‘red’ corresponds to a natural concept – and hence a convex area – of the conceptual space of colour; any colour ‘in-between’ two reds is also red because of the convex geometry.

This thesis proposes that conceptual spaces can serve as an effective framework for the study, design and exploration of generative process. If quality dimensions relevant for a generative process can be identified, then it follows that a generative artist may be able to design into systems and their interfaces the means to navigate the systems’ possibilities in terms of similarity of output. This can afford more effective reasoning about the generative process, as well as allowing for more effective parameter search and discovery. As elusive as quantifiable aesthetic value may be, any such attempts will be best positioned to reach its goals if models that align with the relevant conceptual spaces are used. However judgments of similarity and the identification of natural concepts can only be made if the conceptual space’s quality dimensions are well defined with respect to human perception and knowledge[1].

Within any particular domain, or any one particular generative system, it is not clear what the relevant quality dimensions are. These can be uncovered through empirical study, for instance by collecting many judgments of similarity in experimental contexts[2]. Another approach is to conjecture the relevant quality dimensions around some theoretical frameworks and models of human perception. If the resulting geometrical topologies are such that perceptual similarity follows proximity in the geometrical space, then it would suggest that some aspect of the concep-

[1]Not all representations coincide with conceptual spaces, for instance the computer representation of colour *RGB*, with its red green and blue components, does not coincide with conceptual space of colour, as it does not represent the circular dimension of hue.

[2]This has been done for building the most sophisticated representations of colour, *CIELAB*, where the euclidean distance between points in the space align strongly to similarity judgments.

tual space, for the context at hand, has been uncovered. Further if convex areas in the space have clearly defined characteristics, it is possible that these areas are representing natural concepts.

This bottom-up approach was taken in the design of the *Melody Triangle* interface, the first practical exploration of this thesis.

6.3 The Melody Triangle

The *Melody Triangle* (the subject of chapter 3) is an interface for the discovery of musical content where the parameter space of a stochastic generative musical process, the Markov chain, is 'mapped out' according to the subjective *predicability* of the generated output. It builds on *information dynamics*(Pearce & Wiggins, 2012; Abdallah & Plumbley, 2009; Potter et al., 2007), which provides an account of expectation and surprise in music, and studies these using the tools of information theory (see 2.6.4). Rather than considering music as a static object, information dynamics focuses on subjective expectation and surprise *as they happen*.

The *Melody Triangle* interface defines a mapping between the *quality dimensions* of subjective predictability, and the input parameters to a Markov process. Rather then specifying the parameters of the Markov process directly, the user instead simply specifies a position within a triangle. It is a system for parameter search and discovery, wrapped around the core generative stochastic process of the Markov chain. The Markov chains output streams of symbols, and these are then mapped to musical audio samples.

The triangle maps to an *information space* of predictability, defined by two information theoretic measures, *entropy rate* and *redundancy rate*, these measures are defined in section 3.2. *Entropy rate* represents the average uncertainty about the generated sequence of symbols, whereas *redundancy* quantifies the extent to which the same information is to be found in all parts of the sequence. It was found that these measure correspond to readily perceptual characteristics; high *entropy rate* sequences correspond to sequences with no recognisable structures, and high *redundancy* values are associated with periodic cycles. The triangle's axes align with these two measures, such that a coordinate in the triangle represents particular values for the information measures. When this position is indicated by the user, a Markov chain that has those information properties is used to generate audio. The process of how exactly the audio is generated is detailed in section 3.3.2.

The *Melody Triangle*, with its axes of *entropy rate* and *redundancy*, displays many of the characteristics of a conceptual space. Markov chains spatially close to each other in this space have similar phenomenal characteristics, and certain geometries of the triangle could be said to correspond to 'natural concepts'. The right edge could be said to represent the concept of the 'perfect loop', the top the concept of 'repetition', and the left edge to a concept of 'perfect randomness'. Additionally the left-right direction could be said to correspond to the quality dimension of predictability, and the up-down axis corresponds to the quality dimension of 'many/fewness', as the further down one goes, the more distinct notes appear in the generated pattern on average.

It follows then, that if preference judgments can be gathered from users of the triangle, that it may be possible to begin to associate aesthetic preferences to the conceptual space of predictability. If significant patterns of preferences could be found, then this would give an insight into cross-user objective preferences for predictability in music, with significant potential applications for automatic music generation or evaluation.

6.3.1 The Interactive Installation

Initially the *Melody Triangle* was created as an interactive installation (see 3.4). Visitors in a space were tracked by a *Kinnect* camera, their position in the room mapping to the triangle. They could then by gesturing with their arms, change additional parameters such as instrument choice, volume and notes per beat. This was built as a preliminary exploration and technical demonstration, and although it was considered that it might be of interest to gather preference judgements with the installation, it was found that the situation was far too chaotic to be able to extract any meaningful data.

6.3.2 The Desktop Application

The second incarnation of the triangle was as a desktop application (see 3.5). Here users drag tokens into a triangle on screen with the mouse, and are able to control other parameters with the keyboard. Qualitative feedback from music practitioners was sought, and there were strong indications that *Melody Triangle* could form a useful and effective interface for musical composition or performance (see 3.5.4).

However the main aim of the desktop interface was to attempt to elicit preference judgments from users in a control lab experiment (see 3.5.1). Unfortunately, it was impossible to elicit any such patterns due to a number of flaws in experimental design. Rather than finding musical patterns, it appeared the participants were mainly engaged in an exploratory behaviour aimed at understanding how the system works, and the setup was such that participants would seek to create interesting musical content by continually moving the token rather looking for particular spot in the triangle which would generate musically interesting sequences by itself. It was clear that a different approach was needed, and this motivated the construction of the *Melody Triangle* mobile app.

6.3.3 The Mobile App

The *Melody Triangle* mobile app was developed to attempt to elicit preference judgments from users 'in the wild' (see section 3.6). It was developed for the Android operating system, and is freely available from the Google Play app store[3]. A video demonstration of the app is provided as item *Ex*1 of the illustrative materials accompanying this thesis.

The app allows users to create their own 'compositions' by dragging up to three tokens in and around a triangle on screen, each token representing a melody or 'voice'. They have in the app access to numerous control parameters, including instrument selections for each token (they could choose between piano, drums and bass), they could adjust the register of each instrument and how many notes-per-beat they would generate. Further there were global controls such as scale (they could choose between harmonic, diatonic and pentatonic) and tempo.

Users are encouraged, through notifications in the app, to whenever they find a setting that like, to press a 'like' button. When they do so, the settings of the app would be uploaded to a server for analysis. In this way, the 'liked' settings of users of the app would be crowdsourced. Additionally it is possible for users of the app to listen to the 'liked' settings of other users, this was in the form of 'radio' channel that would select randomly from one of the other users' uploaded settings and play them. The users were encouraged to press the 'like' button if they enjoyed what they heard, or else could press a 'skip' button to hear another song. The idea behind the radio was for it to be a mechanism whereby the users would be able to evaluate each other's songs, providing an extra layer of richness to the collected data. The liked songs were thus used to form a sort of 'chart' of most popular settings. The five most popular songs are provided as item *Ex*2 in the illustrative materials.

[3]https://play.google.com/store/apps/details?id=com.qapps.melodyapp

Implementation Traps

As discussed at length in section 3.6.2, there are a number of difficulties and challenges to developing an audio application on the Android operating system. In particular, the audio libraries that are provided as part of the core Java package have very poor performance. They exhibit unpredictable latency, and are highly inefficient and unstable. Just making the app able to play a note 'in-time' was an excruciatingly difficult task, and having attempted a number of different frameworks and solutions, decent performance was only achieved through some intricate hacks around the Android audio libraries. It is thus highly dissuaded that audio applications are built on the Android platform, iOS is certainly a better alternative.

6.3.4 Results

Was this study thus a 'success', as defined by the criteria defined in 1.2.1, and recapped above? With regards to the first criteria – *correlation between parameter choices and aesthetic value of the output*– the answer is no. As discussed in section 3.6.3, the findings of this practical study were modest. The analysis of the collected data was able to draw mainly trivial observations; users tended to stick to default settings, and they prefer to use extreme settings (maximum BPM, corners of the triangle). Additionally most submissions used all three tokens. With regards to what areas of the triangle were most popular, one thing that was clear: there is a definite preference for the right hand side of the triangle, that is the side with the most predictability. This is perhaps not so surprising as music is, on the whole, structured. There was also an observation that many users placed tokens along the middle of the triangle, the area that would yield sequences that are the most 'informative', with respect to transmitting information that could inform about future events (the area of high *predictive information rate*, see 3.2.4). At first this could have been seen as a significant finding, however it was deemed in the analysis that this preference was most likely a result of interface priming; users placed tokens there because it 'looks nice', similarly the tendency to use corners is likely a result of the fact the triangle *has* corners, and visually would bias token placements towards them. It seems that many of the parameter choices were thus not always related to the aesthetic value of the generated output, but were subject to interface priming. The amount of options available to users were high, and it was unlikely that they had truly explored all the aesthetic options available to them before choosing their 'liked' setting.

With regards to the second criteria – *sufficient volume of data points and significant trends*–

again the answer is no. Although the app did have a large number of installs (3945 as June 17th 2015), there were only 339 submitted entries. The situation with regards to the 'radio' – the chart of most popular settings – the numbers of 'likes' were very small, the most popular song having received only 8 'likes'. This is clearly a feature that did not reach many users.

The number of submitted entries simply was not high enough relative to parameter space at hand, and indicates that there were clearly barriers to user engagement. For a user to do a submission, they would not only need to install the app, but spend long enough with it to understand it, and then feel compelled and motivated enough to submit settings. Each of these steps will have filtered out a large proportion of users.

Even though the aim of the study was mainly to find the areas of the triangle that would prove popular, the users actually had many more parameters to choose from, such as instrument selections, scale, register and tempo. Further the analysis is unable to identify the emergent properties of multiple melodies interacting with each other. There is a difficult balance to strike in such a study, if the amount of features of the app are too limited, and hence the data more clean and parameter space of more manageable dimensions, then it can have an adverse effect on user engagement, as the app would be less 'fun'. Conversely, too many features and the parameter space is too large, and it becomes difficult to extract many significant findings. Additionally as discusses in the analysis, there will have been significant noise introduced with regards to the presentation conditions (varying volumes, headphone models, varying contexts etc), a necessary limitation of many crowdsourcing studies.

This study thus could not be deemed a 'success'; it was not able to yield significant insights of aesthetic preferences with regards to the quality dimensions of predictabillity and surprise.

However there are still some positives to be drawn. The qualitative feedback from the users of the desktop application showed that they were readily able to identify the properties of the different areas of the triangle, and did recognise the potential of this interface toward musical creation. The *Melody Triangle* demonstrates how it is possible to demarcate the quality dimensions of a generative process, and with these quality dimensions, to create a useful and usable interface for the explorations of the creative possibilities of a generative process.

6.3.5 Future Work

In section 3.7, a number of avenues for future work with the *Melody Triangle* were discussed. The first is to explore ways to tackle the issue of user engagement. It is suggested that this could be achieved by making the app closer to a professional musical interface. This could be done offering fine-grains controlled over symbol to note mappings, and provide features that would allow it be used professionally in a performative or recording context, such as the ability to output MIDI or OSC to integrate the app with other studio tools. If the app became a genuinely useful musical tool, then it would be more likely to be able to reach a critical mass of use, and then perhaps enough data could be gathered.

Another avenue of future work involves developing a more sophisticated 'listener model' to the triangle. Currently the information measures are based around first order Markov processes, and thus are not able to represent long term musical structures. A clear improvement would be to use generative processes with long term musical structure. This would require further research to identify the information measures that relate to the quality dimensions of predictability for these more sophisticated generative processes.

The final area of suggested future works involves exploring the possibility that the *Melody Triangle* affords other domains. The Markov chains output abstract symbols, and these could be mapped to anything, not just musical notes. They could even be applied to sequences of lights or animations. A particularly exciting potential area of exploration is to see what such an interface could provide in the timbral domain, perhaps as a controller for granular synthesis.

6.4 Keyebernates

Keyebernates is the second practical exploration of this research, the subject of chapter 4. It was developed in part as a reaction to the difficulties faced in the *Melody Triangle* study, in particular the difficulties that interface priming brings to eliciting aesthetic judgments. Is it possible to have 'no interface', and make the very act of consuming the artefact be the means by which the possibilities of a generative process is explored, and judgments of value extracted?

Keyebernates is, like the other practical explorations of this research, a system for parameter search and discovery. The generative process it controls is a trivially simple one; white adjacent circles that overlap to make patterns on a black background. In stark contrast to the *Melody Triangle*, the core generative process had just three parameters; circle size, and two 'offset' parameters that determined in what way the circles would overlap. The mechanism for navigation and control was through eye-tracking. *Keyebernates* places its core simple generative process in an 'ecosystemic' installation(see 2.2.4). It sets up a continuous feedback loop between generated output and the viewer, via their gaze. This cybernetic 'structural coupling' between the observer and the system yields an *emergent* interaction, manifesting itself as the navigation of the parameter space of the generative process. A video illustrating the dynamics of *Keyebernates* is provided in the illustrative materials accompanying this thesis as item *Ex*3.

It can not quite be said that *Keyebernates* had 'no interface' – eye trackers are an interface after all – but the setup of the system was such that there would be as little a distinction between the act of contemplating the work and controlling it, and indeed a key element was that contemplators do not know that they are affecting the output.

Keyebernates was designed in part as an approach to dealing with the discontinuities inherent in generative design; unlike traditional non-generative methods of creation, where there is an immediate and clear causality between the artists' action and its consequence, in generative design there is often a discontinuity and lack of embodied control as the designer choses parameters, awaits its consequence, and then evaluates the outcome. By using eye-tracking and gaze points as the modality of input, the viewer navigate the parameter-space of the generative system with their eyes, the parameters most gaze upon being reinforced, in a continuous and reactive navigation of parameter and solution space.

A key element of *Keyebernates* is that when a viewer observes it, they do not directly see changes as they happen. Instead by using the real-time data from the eye-tracker, *Keyebernates* ensures that what is attended to does not change, and rather all change happens slowly and subtly in ares of the screen not currently attended to.

The navigation was engendered by a counter balancing two forces (see 4.2); a noise source that increases variety in the parameter values, a process of *(*differentiation)[4]. Conversely the feedback from the eye tracker drives a process of *integration*[5], where the parametric variety of the generated output increases. This delicate balancing and tuning of the integrative(gaze) and differentiating(noise) forces, enable the system *as a whole*, together with its perceiver, to navigate state space.

Two experiments were carried out with *Keyebernates* (see 4.3), in both experiments, the participants were only told that their gaze had any impact on the output till afterwards, and instead they were lightly deceived and told that they would be watching a video. The first experiments aligns with the main goals of this research, as it attempts to elicit a relationship between the pa-

[4]In the non-mathematical meaning of the word; to increase differences

[5]Again in the non-mathematical meaning of the word; to make things more similar

rameter space of a generative process and the perceived aesthetic value. This was done simply by placing participants in front of *Keyebernates* and having them observe the screen, their passive, *non-volitional* gaze driving the direction of parameter space navigation. The hypothesis was that gaze would linger more on the 'interesting' outputs, as some of the research surveyed seemed to indicate (see 4.1.1), and as such *Keyebernates* would guide the navigation towards the more 'interesting' areas of parameter space.

The second experiment tested how the system would react to guided *volitional* gaze. Here the participants were shown a set of patterns on a piece of paper, and were asked to look for each of those patterns in the 'video', and that if it appears, to try to spot it as soon as possible. This was to test if the system could be used to 'will' a navigation towards a specific area of parameter space, and thus give insight into the potential practical application of *Keyebernates* as a design interface.

6.4.1 Results

Only the first experiment, navigation with respect to passive *non-volitional* gaze, is compatible with the criteria for success defined in 1.2.1 and recapped above.

In the first experiment the analysis of the navigation of users under passive gaze, all users pulled the system towards circles of increasing size. This indicated that gaze was clearly attracted towards images of greater visual complexity, as the bigger the circles, the more they would overlap with each other. Navigation with respect to the other parameters (of *offset in x* and *offset in y*) varied for each participant.

However a key difficulty of this study was that the *Keyebernates* caused significant fatigue to the participants, so much so that some did not feel comfortable carrying out the whole study. This was likely because *Keyebernates* is un-natural in that it is *gaze-contingent*; it behaves and appears differently with respect to how it is gazed upon. The viewer's visual system is continually trying to build a coherent representation of the scene, while the scene is shifting constantly, but only in the peripheral vision. It is thus perhaps not surprising that the participants found interfacing with it fatiguing.

Was it thus successful? With regards to the first criteria for success – *correlation between parameter choices and aesthetic value of the output*– it was not. It is clear that for all subjects, passive gaze pulled the system towards larger circles and thus more intricate patterns. However that does not mean that patterns with larger circles are more beautiful, only that they are, generally, more 'eye-catching' than patterns in neighbouring areas of parameter space. It is just as possible that gaze was attracted to the 'complexity', or simply attracted to brightness.

With regards to the second success criteria – *sufficient volume of data points and significant trends*– again the answer is no. The slowness of the navigation of through parameter space made it such only a small portion of the parameter space would be explored, and further the fatigue that it induced made it such that prolonged coupling to *Keyebernates* could not be carried out, and greatly limited the amount of time participants could interface with the system, and hence the amount of data that could be gathered. As such *Keyebernates* could not be deemed a 'success'.

The second experiment – the search guided by volitional gaze – tested its capacity as a potential control interface for generative designs. In this respect it had mixed results. The target patterns that the participants 'looked for', were of varying parametric distance from the starting position in parameter space. All of the participants were able to pull the system towards the target pattern that was 'near' to the system's starting position. But only some of them could pull the system to the pattern that was at an intermediate distance, and none of the participants pulled the system towards the 'far' search target.

It seems that the one possible difficulty relates to the fact that parametric similarity does not coincide with phenomenal similarity, even for such a trivial generative process as these overlapping circles. For the 'far' search target, some participants would guide the system in the completely wrong direction, and at other times would get stuck in a certain areas of parameter space. Another problem was the speed of navigation, as *Keyebernates* relies on a slow navigation – as otherwise changes in the peripheral vision would attracted the gaze by virtue of the change – it would take an impractical amount of time to make the system travel far in parameter space. *Keyebernates* does not scale well, even though it could also be used with complex generative visual processes of many parameters, because there is a constant and limited amount of screen real-estate, the more parameters there are the slower the navigation becomes.

Keyebernates, in its current form, is of limited practical application. However as discussed in section 4.4.3, it has value as an art-work in it own right.

6.4.2 Future Work

A number of areas were identified for future work with *Keyebernates* (see 4.4.2). The first is to find means of mitigating fatigue. This could be done by employing higher resolution displays, as this would make changes in the peripheral vision smoother and less noticable, and hence may make interacting with the system less tiresome. Another approach is to slow down the speed of navigation further to make changes in peripheral vision even more subtle. Although this would limit its practicality further, other avenues can be explored to make the exploration of parameters faster.

A key area of further research is to find a way to re-align the parameters through an intermediate mapping layer, in such a way that parametric similarity *does* coincide with perceptual similarity. If such a mapping where applied to *Keyebernates*, it could make navigating the possibilities of the designs much more effective. In much the same way that the *Melody Triangle* re-aligns the parameters of the Markov chain to the quality dimensions of predictability, the same could be done of visual generative processes. It is unclear what the relevant quality dimensions are for any particular generative process, but it is possible that such a mapping could be uncovered by collecting many judgments of *similarity* from crowds.

If it is to become a practical design tool, then eye-tracking should not be the sole mechanism of control; *Keyebernates* is only really effective at engendering a detailed exploration of a small area of parameter space. Parameters could be controlled by some other mechanism, such as a physical haptic controller, in addition to the gaze-tracking. The designer could then navigate towards an approximate design with their hands, and then switch one or more parameters to be under gaze control to navigate through a detailed area of parameter space.

A contribution of this thesis was to demonstrated that with *Keyebernates*, volitional-gaze could be used to navigate the parameter-space of generative designs across modest parametric distances. Although not practical in its current form, this nevertheless suggests that such an approach could be used in the design of an effective interface for parameter discovery.

6.5 EvoColour

The third and final practical study of this research was *EvoColour*, the subject of chapter 5.

It was motivated largely by the lack of ‘success’ of the previous two studies, and as such its design drew on lessons learned from these studies. It was ensured that the parameter space of the generative process to be explored was not exceedingly large, and that there would be a reliable way of extracting judgments of value from the participants.

EvoColour is a crowdsourced, interactive evolutionary system where populations of simple images ‘evolve’ in response to preference selections made by members of the public over the internet. Like the other studies of this thesis, *EvoColour* 'wraps around' a core simple generative process, and is also a system for parameter search and discovery. The core generative process is the *Markov image* (see 5.4.1) , and is realised as a sequence of fifty concentric circles of up to three colours (the *phenotype*), with its parameters consisting of a Markov chain that outputs a sequence of symbols, and three colour values that map to these symbols (the *genotype*).

Using a genetic algorithm inspired by Darwinian evolution, images that are more popular – as determined by preference judgments of users – are selected for survival, while those that are least poplar are 'killed off'. The popular images then sexually reproduce to fill the slots vacated by the killed-off images (see 5.4.2). The aesthetic judgments consisted of users indicating which of two randomly selected images from the population they preferred, additionally the viewers could indicate that they had 'no preference'.

When the 'fittest' Markov images in a population sexually reproduce (see 5.4.3), characteristics of the parents are passed to the children. These include the relationship between the colours in colour space, and the properties of the sequence of the colours, as encoded in the Markov chains. The calculations on colour space were carried out with respect to the *CIELAB* colour model (see 5.2), as it corresponds closely to the *conceptual space* of the human perception of colour. Each generation, images would have a chance of 'mutation', whereby some of the parameters of the image would be randomly changed. This ensured that the populations would not homogenise too quickly and prevent the evolution from getting stuck at a 'local maxima' of fitness.

Multiple evolutions on populations of 500 images were carried out, and additionally there was a 'control population' that would not evolve, but its images were ranked according to preference. The images of the evolved populations, as well as that of the control population, are available as items $Ex4-7$ of the illustrative materials accompanying this thesis.

6.5.1 Results

Extensive numerical analysis was carried out on the evolved populations (see 5.6), and a large number of observations could be derived. The observations, summarised in table 5.13, included identifying trends in preference with regards to overall colour, colour combinations, as well as the arrangements and relative amounts of theses colours in an image.

There is evidence of a cross-user *consensus*, of a global agreement over which images are to be preferred over others. This is evident even just by looking in the control population, *CP*, where it was clear that some images are consistently selected over others. Some of these observations support previous work in colour research (see 5.3). For instance the well known observations that, overall, blue is the 'world's favourite colour', and that yellows are disliked, were echoed in the analysis of the data collected by *EvoColour*. Additionally some less well known patterns in colour combination preferences were also reflected, for instance the observation that high contrast in lightness in an image is generally deemed desirable.

However many new observations of aesthetic preference not previously identified in the literature of colour research were brought to light with *EvoColour*, in particular, observations relating to the *arrangements* of colours in an image, as well as the *relative amounts* of these colours. This includes the observations that both images that embody perfect repetition and those that lack all structure were disfavoured over images of intermediate regularity, that sparsity is favourable, and that images with unequal amounts of colour are preferred. Additionally it was shown that colour preferences are highly contingent on how these colours are arranged (images of high entropy had differing overall colour preferences to images of low entropy), an area not previously explored in the literature.

Was *EvoColour* thus a 'success', as defined by the criteria defined in 1.2.1, and recapped above? With regards to the first criteria – *correlation between parameter choices and aesthetic value of the output*– one needs to determine two things; did the selections really correspond to viewers notion of the aesthetic value of the images? And did the evolutionary processes of *EvoColour* drive the parameters for populations of Markov images in way to correlated with the aesthetic preferences of the participants?

The context of contemplation in *EvoColour* shows two images as pairs; the situation is explicitly setup for viewers to indicate a preference. As such the selections are indicators of *relative* aesthetic value: the viewer does not communicate wether or how much they like an image, rather they indicate just wether they like one image more than the other. Such comparisons distill the nuanced and subtle process of aesthetic contemplation to a binary selection. It was found in the analysis of the non-evolving control population, that there is evidence of cross-user *consensus* over which images are to be preferred over others. This is despite some the variability in presentation conditions that such a crowdsourcing project necessarily engenders. That there exists such a consensus, and given that aesthetic judgments do exhibit this kind of consensus(McMahon, 2011), it would seem that the selections did align with an objective dimension of aesthetic value.

With regards to wether the evolutions did really make images 'fitter', and drive the parameters of the population images in a way that correlated with aesthetic preference, this was verified empirically by making having images from the first and last generations be subject to preference judgments (see 5.6.2), the mean fitness of the populations increased across the generations. As such the first criteria for 'success' could be said to have been met.

With regards to the second criteria – *sufficient volume of data points and significant trends*– it was clear that the analysis of *EvoColour* succeeded in making a large number of observations. There were initial concern that the Markov image would have either too great a potential complexity for the evolution to lead to identifiable trends with respect to the number of selections that could be gathered, or that it would be too difficult to learn anything new in analysis. It seems the balance was about right; the state space of the Markov image is suitable for evolutionary processes, while simultaneously lending itself well to analysis. The ability to draw so many observations suggest there were enough data points, fulfilling criteria $C2$.

As it would appear that both criteria $C1$ and $C2$ were met, it seems safe to suggest that *EvoColour* was a 'success'; It was able to uncover objective, cross-user elements of aesthetic value with respect to colour and colour combinations in Markov images.

6.5.2 Future Work

A number of areas of future work can be explored with *EvoColour* (see 5.7.2). There are for instance a number of potential practical applications to the findings of *EvoColour*. An exciting potential application is to frame the observations as design heuristics, and to incorporate them into the construction of smart design tools. Such tools could provide an objective aesthetic evaluation of an artefact under construction, providing suggestions for ways to improve the colour aspects of the designs. It is even conceivable these heuristics could inform the construction of software that could stand in and autonomously make creative decisions that are likely to appeal to many.

Another area of further work is to explore various ways of improving *EvoColour*. One of these is to seek ways to pull out families of interesting parameter values, rather than have the evolution go towards some form of global average. One approach could be to curate the starting population of images. For instance, if one is out to find the parameters that would yield the most popular Markov images that are mostly red, the population in the first generation could be seeded with exclusively red images. An alternative would be to prevent some attribute of the images from being affected by the evolutionary process. For instance preventing the blue dimensions of the colours from being changed in the mating or mutating process, would disable the tendency of populations of converging towards the blue areas of colour space, and thereby allowing other attributes of these images to come to the for, revealing differing areas of the fitness landscape.

All evolutionary algorithms will have their own set of imperfections that 'bias' the evolution in one direction of state space over others. An area of further work is to improve the evolutionary and mating algorithms of *EvoColour* to be more resistant to bias, and carry through the aspects of the images that correspond to aesthetic preference more efficiently across the generations. Additionally the way the analysis was carried out could be improved by identifying more features for the detection of trends and patterns.

Each observation that could be made as a result of the analysis can be interpreted as a design heuristic. A final intriguing area of further research is to explore to possibility of using these heuristics to define computational fitness functions for the automatic synthesis of images of high fitness. How much importance should be placed on each heuristic difficult to determine. It would be of interest to develop a kind of meta-evolution, where images evolved with differing heuristic weightings are then subject to preference judgments from human contemplators. This could help uncover which heuristics are most important in determining aesthetic value; for instance is overall blueness of an image more or less important then having a large amount of lightness contrast in the image? If a set of weightings for these heuristics could be determined, it is conceivable that it may be possible to arrive at something close to an 'archetypal image'; the image that rests at the hypothetical highest peak of fitness.

The fitness landscape, as we know, will vary from person to person, as well as vary with other factors such as differing presentation conditions. But as there clearly are agreements across the judgments of individuals, then it implies that there is also an 'average' fitness landscape; the landscape that represents only the objective dimension of human aesthetic value. In a sense, all the studies of this thesis were attempts at uncovering this average landscape, and to look for its highest peak.

6.6 Reflections

Driven by a fascination for the creative possibilities of algorithmic thought, this research set out to elucidate its nature. In delving into the past, it was recognised that the use of such processes could indeed be found everywhere. But what motivates artists and musicians to employ these processes?

Setting up these experiments brought about a familiar sense of expectancy and excitement that artists who use generative processes in their work no doubt will recognise; the mysterious and seductive anticipation of setting up an algorithmic situation, and awaiting to see the emergent result.

It was not just scientific curiosity that motivated their inception, but also an artistic curiosity. It was this curiosity that initiated the search for the archetypal image in *EvoColour*, that wondered what patterns would emerge from the ecosystemic installation that is *Keyebernates*, and that sought to find alluring minimalist musical patterns through the crowdsourcing of the *Melody Triangle*.

The three experiments of this thesis are not just works of science, but can be viewed as artistic generative compositions in and of themselves. When considering these systems, *together* with their users or viewers, as higher-level generative processes running *in time*, they exhibit properties of complex, emergent self-organising systems[6]. As the database of collected songs in the *Melody Triangle* continue to build up, as a viewer coupled to *Keyebernates* navigates parameter space, and as the populations of images in *EvoColour* continue to be subject to the judgments of viewers world-wide, a self-organisation occurs as a process of the data-collection. The three experiments were setup such that aesthetic preference judgments – captured through parameter selections – was an organisational force. The analyses then sought to find the ways to best account for this self-organisation.

It is difficult to claim an absolute creative responsibility for the works of this thesis. On the one hand, these are clearly the work of the present author, who designed the generative processes and the elements that composed the final forms that were generated. On the other hand, these works require the input of participants and observers for the selection of parameters, and hence to exist at all. As amorphous entities, with distributed authorship and agency, they are perhaps

[6]Self-organisation can be understood as a ‘reduction of entropy’, or rather, that self-organising systems extract order from their ‘environments’ (see 2.5). In these practical studies, the ‘environment’ is the humans interacting with the systems.

effectively characterised as 'art-systems', where by the traditional clearly defined roles of artist, artwork and audience no longer apply.

The scientific curiosity that drove these experiments aimed to uncover insights into the nature of aesthetic judgments; to elucidate the relationship between their objective and subjective dimensions. There is a fundamental paradox here. If, say, a preference for blueness is normative, then it is objective and must have a factual basis. Yet a preference for blue is a value that expresses an attitude. Aesthetic judgments may be unique in their nature of being both expressions of subjective value, while having a link to an objective truth.

These experiments collected many small decisions from numerous people and applied techniques of scientific analysis to uncover the latent commonalities between aesthetic opinions. With these commonalities a map relating the objective aesthetic judgments to the generative process could be roughly sketched. This map removes the subjective idiosyncrasies of taste, revealing the geography of a landscape that is common to us all.

The map that *EvoColour* revealed, as represented in the heuristics that the analysis uncovered, is tied to the specific generative process with which it was engendered. Yet it represents a kind of aesthetic consensus which could be used to automate the creation of art, but it would be art with no artist. Even if the present author is responsible for sketching it, the map does not 'belong' to anyone; it is a projection – a shadow – of our common cognitive mechanisms.

Seeking out this objectivity in aesthetic value is a way of probing the underlying mechanisms of perception, to uncover what as humans we all share. Science is rapidly mapping out the perceptual building blocks that make up our sensations; the quality dimensions and conceptual spaces through which we make sense of the world. But even though science is increasingly yielding insights into the mechanisms of aesthetic appreciation, it is a long way from having a complete picture.

Perhaps this is a good thing. Perhaps art should remain a mystery, and may its capacity to move us always be just beyond the reach of scientific comprehension.

Appendix A

EvoColour - EP3 list of features

Table A.1: Features of images in $EP3$. $\mu gen1$ is the mean of the values in generation 1, $\mu gen82$ is the mean of the values in generation 82, $\Delta\mu$ is the change between the two means. $\sigma gen1$ is the standard deviation of generation1, $\sigma gen82$ is the standard deviation at the final generation. $\Delta\sigma$ represents the *homogenisation*, which is the relative magnitude of the decrease in standard deviation between the fist and the last generation. Table is sorted by decreasing *homogenisation*.

Feature	$\mu gen1$	$\mu gen82$	$\Delta\mu$	$\sigma gen1$	$\sigma gen82$	$\Delta\sigma$
$I_{Y;X}$	0.30	0.17	-0.12	0.27	0.16	0.43
$\max\lvert\vec{c}\rvert$	9.17	12.47	3.30	11.81	6.92	0.41
$\min\Delta b^*$	28.07	17.96	-10.11	25.24	14.99	0.41
$\max(a^*)_{2/3}$	33.29	32.67	-0.62	35.14	22.21	0.37
$R(b^*)_{2/3}$	63.13	43.76	-19.37	41.16	26.54	0.36
$\sigma(b^*)$	27.75	16.30	-11.45	15.41	9.95	0.35
$\max(a^*)_{3/3}$	32.25	31.26	-0.99	35.66	23.20	0.35
$\mu\Delta b^*$	52.03	33.82	-18.22	28.96	19.02	0.34
$\lvert\Delta c\rvert$	27.84	17.60	-10.24	12.85	8.50	0.34
$R(b^*)_{3/3}$	61.85	42.55	-19.30	42.03	28.53	0.32
$\sigma(a^*)$	25.35	16.74	-8.60	14.89	10.11	0.32
$R(a^*)_{2/3}$	56.16	44.45	-11.72	38.44	26.55	0.31
$\lvert C\rvert$	2.82	2.92	0.10	0.39	0.27	0.31
$\min\Delta a^*$	23.14	17.79	-5.35	23.23	16.08	0.31
$\mu\Delta a^*$	46.49	34.74	-11.75	27.69	19.52	0.29
$\max\Delta b^*$	72.33	48.12	-24.21	36.75	26.13	0.29
$R(a^*)_{3/3}$	56.37	42.40	-13.97	38.57	27.43	0.29
$R(b^*)$	73.36	48.80	-24.56	36.93	26.31	0.29
$H^2_{Y\|X}$	0.63	0.67	0.04	0.31	0.22	0.28
$H^3_{Y\|X}$	0.44	0.52	0.08	0.22	0.16	0.28
$\max(a^*)$	38.48	34.65	-3.83	30.81	22.48	0.27
$R(a^*)$	65.67	49.35	-16.31	34.13	25.15	0.26
$\max\Delta a^*$	65.04	48.51	-16.53	33.83	25.01	0.26
$I^2_{Y;X}$	0.56	0.36	-0.20	0.31	0.23	0.26
$\min\Delta h^\circ$	48.29	33.20	-15.09	42.94	31.86	0.26
$\max(a^*)_{1/3}$	35.88	30.94	-4.94	33.18	24.82	0.25
$R(b^*)_{1/3}$	68.45	40.97	-27.48	38.69	29.07	0.25
$\mu(a^*)_{2/3}$	4.48	10.30	5.82	31.21	23.61	0.24

Feature	$\mu gen1$	$\mu gen82$	$\Delta\mu$	$\sigma gen1$	$\sigma gen82$	$\Delta\sigma$
$\mu(a^*)_{3/3}$	4.00	10.01	6.01	30.78	23.30	0.24
$\min(a^*)_{2/3}$	-22.87	-11.77	11.10	32.53	25.01	0.23
$\mu(a^*)$	4.48	10.44	5.96	29.80	22.92	0.23
$\mu(a^*)_{1/3}$	4.93	10.98	6.05	30.21	23.70	0.22
$R(a^*)_{1/3}$	60.41	41.81	-18.60	35.97	28.28	0.21
$\min(C^*)_{2/3}$	38.08	27.74	-10.33	20.15	15.89	0.21
$\min(b^*)_{2/3}$	-27.42	-39.74	-12.32	36.04	28.65	0.21
$\min(a^*)$	-27.19	-14.70	12.49	28.84	23.27	0.19
$\min(a^*)_{3/3}$	-24.12	-11.14	12.98	31.68	25.63	0.19
$\min(b^*)_{3/3}$	-26.67	-39.26	-12.59	36.32	29.93	0.18
$\min(a^*)_{1/3}$	-24.53	-10.87	13.66	31.14	25.87	0.17
$\min(C^*)_{3/3}$	38.00	28.32	-9.68	20.02	16.67	0.17
$\max(h^\circ)$	273.80	287.52	13.73	73.68	61.74	0.16
$H_{Y\|X}$	0.90	0.86	-0.04	0.41	0.35	0.15
$\min(b^*)_{1/3}$	-30.62	-39.18	-8.56	33.11	28.64	0.13
$\max(b^*)_{3/3}$	35.18	3.29	-31.90	34.16	29.71	0.13
$\mu(b^*)_{3/3}$	4.71	-19.40	-24.11	31.32	27.36	0.13
$\max(b^*)_{2/3}$	35.70	4.01	-31.69	33.63	29.48	0.12
$\min\Delta C^*$	14.89	13.23	-1.65	14.09	12.49	0.11
$\min(C^*)$	35.19	26.27	-8.92	17.29	15.38	0.11
$\mu(b^*)_{2/3}$	5.00	-19.13	-24.13	31.48	28.02	0.11
$R(C^*)_{2/3}$	34.74	34.21	-0.52	23.80	21.23	0.11
$\sigma(C^*)$	15.45	12.91	-2.54	9.49	8.47	0.11
$\mu(h^\circ)_{3/3}$	183.66	239.44	55.78	108.14	97.24	0.10
$\min(b^*)$	-33.02	-41.82	-8.80	31.23	28.20	0.10
$\mu(b^*)$	4.75	-19.56	-24.31	30.16	27.27	0.10
$\mu(h^\circ)_{1/3}$	185.95	237.29	51.35	109.12	98.68	0.10
$\min(C^*)_{1/3}$	36.87	29.16	-7.71	18.80	17.15	0.09
$\mu(h^\circ)$	183.62	236.16	52.54	108.72	100.13	0.08
$\mu(b^*)_{1/3}$	4.55	-20.13	-24.68	30.16	27.81	0.08
$\mu(h^\circ)_{2/3}$	182.52	231.92	49.40	109.66	101.22	0.08
$R(C^*)_{3/3}$	34.21	33.01	-1.20	23.97	22.35	0.07
$\mu\Delta h^\circ$	89.41	67.43	-21.98	40.39	38.05	0.06
$\mu\Delta C^*$	28.16	26.37	-1.78	16.98	16.00	0.06
$R(C^*)$	40.17	38.23	-1.94	21.63	20.44	0.06
$\max\Delta C^*$	39.56	37.87	-1.70	21.57	20.42	0.05
$\mu(C^*)_{3/3}$	54.89	43.91	-10.98	18.55	17.61	0.05
$\min\Delta(\Delta E^*)$	30.35	27.75	-2.60	15.87	15.13	0.05
$\max(b^*)_{1/3}$	37.83	1.79	-36.05	31.75	30.61	0.04
H_X	1.20	1.04	-0.17	0.37	0.36	0.03
$\mu(C^*)_{2/3}$	55.13	44.03	-11.10	18.26	17.87	0.02
$\mu(C^*)$	55.11	44.23	-10.88	17.64	17.28	0.02
$I^3_{Y;X}$	0.75	0.51	-0.23	0.36	0.35	0.01
$\max(b^*)$	40.33	6.98	-33.35	30.09	29.77	0.01
$\mu(C^*)_{3/3}$	72.21	61.33	-10.88	21.57	21.40	0.01
$\max(C^*)_{2/3}$	72.81	61.96	-10.86	21.07	21.00	0.00
$R(C^*)_{1/3}$	37.18	31.74	-5.44	22.90	22.96	-0.00
$\mu(C^*)_{1/3}$	55.29	44.73	-10.56	17.65	18.10	-0.02
$\max(C^*)$	75.36	64.50	-10.86	19.44	20.00	-0.03
$\max\Delta(\Delta E^*)$	64.34	64.02	-0.32	18.55	19.16	-0.03
$\min(L^*)$	40.25	22.76	-17.50	16.91	17.69	-0.04
$\min(L^*)_{3/3}$	43.37	25.19	-18.18	18.87	19.94	-0.05
$\max\Delta h^\circ$	121.19	95.15	-26.03	46.23	49.32	-0.06
$\min(L^*)_{2/3}$	42.95	24.32	-18.63	18.30	19.64	-0.07
$\max(C^*)_{1/3}$	74.05	60.89	-13.15	20.22	21.77	-0.07
$\mu\Delta(\Delta E^*)$	47.87	48.70	0.83	15.34	16.53	-0.07

Feature	$\mu gen1$	$\mu gen82$	$\Delta\mu$	$\sigma gen1$	$\sigma gen82$	$\Delta\sigma$
$\sigma(L^*)$	12.85	17.44	4.58	7.50	8.41	-0.11
$\max\Delta L^*$	32.91	48.96	16.05	17.39	19.58	-0.11
$R(L^*)$	33.49	49.88	16.39	17.49	19.80	-0.12
$R(L^*)_{2/3}$	28.75	45.90	17.14	19.18	21.93	-0.13
$\min\Delta L^*$	13.26	16.58	3.32	12.90	15.10	-0.15
$\min(L^*)_{1/3}$	41.45	25.86	-15.59	17.52	20.60	-0.15
$\mu\Delta L^*$	24.11	36.22	12.11	13.90	16.42	-0.15
$R(L^*)_{3/3}$	28.74	44.04	15.31	19.48	23.49	-0.17
$R(L^*)_{1/3}$	31.12	41.98	10.85	18.23	24.20	-0.25
$\max(L^*)_{2/3}$	71.70	70.21	-1.49	15.80	21.31	-0.26
$\mu(L^*)_{3/3}$	57.59	43.96	-13.63	15.61	21.20	-0.26
$\max(L^*)$	73.75	72.64	-1.11	13.54	18.87	-0.28
$\mu(L^*)_{2/3}$	57.58	43.85	-13.73	15.61	21.89	-0.29
$\mu(L^*)$	57.44	43.71	-13.73	14.90	20.94	-0.29
$\max(L^*)_{3/3}$	72.11	69.24	-2.87	15.59	22.48	-0.31
$\mu(L^*)_{1/3}$	57.16	43.33	-13.82	14.80	21.59	-0.31
$\max(L^*)_{1/3}$	72.57	67.84	-4.74	14.70	22.20	-0.34

Appendix B

Illustrative Materials

One data CD is provided as illustrative material for this thesis.

It has a folder for each of the three studies of this thesis: the *Melody Triangle*, *Keyebernates* and *EvoColour*. The contents of each of the folders have also been archived online. The paths and URLs to these materials are provided in table B.1.

The Melody Triangle

The illustrative materials for the *Melody Triangle* consist of a demo video of the mobile app, and the top 5 songs of the *Melody Triangle Radio* as of June 17th 2015.

Keyebernates

The illustrative materials for *Keyebernates* consist of a video showing the dynamics of the system.

EvoColour

The illustrative materials for *EvoColour* consist of the all the images of the control population *CP*, sorted in order of popularity. Additionally all the images from the final generations of $EP1$, $EP2$ and $EP3$ are provided.

Table B.1: Illustrative Materials

Ex #	Path / URL	Description
/MelodyTriangle https://youtu.be/U7w69aabkqk		
$Ex1$	/MelodyTriangle_DemoVideo.mov	Demo video of the Melody Triangle app
$Ex2$	/MTTopTrack1-5.aif	Top five tracks of the Melody Triangle Radio

/Keyebernates https://youtu.be/1-NSAf7CGoA		
$Ex3$	/Keyerbenates_WalkThrough.mov	A video showing the dynamics of Keyebernates

/EvoColour https://www.flickr.com/photos/140600562@N05/collections/72157663025467944/		
$Ex4$	/CP/CP_x.png	All 1000 images of control population CP. Sorted in order of popularity.
$Ex5$	/EP1/EP1_x.png	All 500 images of generation 86 of EP1
$Ex6$	/EP2/EP2_x.png	All 500 images of generation 85 of EP2
$Ex7$	/EP3/EP3_x.png	All 500 images of generation 81 of EP3

References

Abdallah, S., & Plumbley, M. (2009). Information dynamics: patterns of expectation and surprise in the perception of music. *Connection Science*, *21*(2-3), 89–117.

Abdallah, S., & Plumbley, M. (2010). Predictive information, multi-information, and binding information. *Technical Report C4DMTR-10-10, Queen Mary University of London*.

Acharya, P. K. (1946). *An encyclopaedia of Hindu architecture* (Vol. 7). Munshiram Manoharlal Publishers Pvt. Limited.

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science-AAAS-Weekly Paper Edition-including Guide to Scientific Information*, *262*(5142), 2042–2044.

af Klint, H., Kunz, E., & Martin, A. (2005). *3 X Abstraction: New methods of drawing* (M. C. de Zegher & H. Teicher, Eds.). Yale University Press.

Allan, M., & Williams, C. K. (2005). Harmonising chorales by probabilistic inference. *Advances in neural information processing systems*, *17*, 25–32.

Alloway, L. (1966). *Systemic painting*. Solomon R. Guggenheim Foundation.

Ariza, C. (2005a). Navigating the landscape of computer-aided algorithmic composition systems: A definition, seven descriptors, and a lexicon of systems and research. In *Proceedings of the 2005 international computer music conference* (pp. 765–772).

Ariza, C. (2005b). *An open design for computer-aided algorithmic music composition*. Universal Publishers.

Arnheim, R. (1954). *Art and visual perception: A psychology of the creative eye*. University of California Press.

Ascott, R. (1990). Is there love in the telematic embrace? *Art Journal*, *49*(3), 241–247.

Ashby, W. R. (1991). Principles of the self-organizing system. In *Facets of systems science* (pp. 521–536). Springer.

Atlan, H. (1979). *Entre le cristal et la fumée: essai sur l'organisation du vivant*. Éditions du Seuil.

Back, T., Fogel, D. B., & Michalewicz, Z. (1997). *Handbook of evolutionary computation*. IOP

Publishing Ltd.

Beaumont, J. G. (n.d.). Lateral organization and aesthetic preference: The importance of peripheral visual asymmetries. *Neuropsychologia*, *23*(1), 103–113.

Beddard, H. (2009). Computer art at the v&a. *V&A Online Journal*.

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics*, *8*(5), 279–286.

Biles, J. (1994). Genjam: A genetic algorithm for generating jazz solos. In *Proceedings of the international computer music conference* (pp. 131–131).

Birkhoff, G. D. (1933). *Aesthetic measure*. Harvard University Press.

Blackwell, T. (2008). Swarm granulation. In *The art of artificial evolution* (pp. 103–122). Springer.

Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. Routledge.

Boden, M. A., & Edmonds, E. A. (2009). What is generative art? *Digital Creativity*, *20*(1-2), 21–46.

Brown, P. (2001). Stepping stones in the mist. In *Creative evolutionary systems* (pp. 387–407).

Bulling, A., Roggen, D., & Troster, G. (2011). What's in the eyes for context-awareness? *Pervasive Computing, IEEE*, *10*(2), 48–57.

Buswell, G. T. (1935). *How people look at pictures: A study of the psychology of perception in art*. University of Chicago Press.

Cage, J. (1970). Extract from 'A year from Monday'. In J. Reichardt (Ed.), *Cybernetic serendipity, the computer and the arts.* Studio International Special Issue.

Candy, L., & Edmonds, E. (2002). *Explorations in art and technology*. Springer.

Cariani, P. (1993). To evolve an ear. Epistemological implications of Gordon Pask's electrochemical devices. *Systems Research*, *10*(3), 19–33.

Cavanagh, P. (2011). Visual cognition. *Vision Research*.

Chevreul, M. E. (1855). *The principles of harmony and contrast of colours, and their applications to the arts*. Longman, Brown, Green, and Longmans.

Chowning, J. M. (1973). The synthesis of complex audio spectra by means of frequency modulation. *Journal of the Audio Engineering Society*, *21*(7), 526–534.

Cohen, H. (1995). The further exploits of AARON, painter. *Stanford Humanities Review*, *4*(2), 141–158.

Cohen, J. E. (1962). Information theory and music. *Behavioral Science*, *7*(2), 137–163.

Collins, N., & McLean, A. (2014). Algorave: A survey of the history, aesthetics and technology of live performance of algorithmic electronic dance music. In *Proceedings of the international conference on new interfaces for musical expression.*

Conklin, D., & Witten, I. H. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research*, *24*(1), 51–73.

Conway, J. (1970). The game of life. *Scientific American*, *223*(4), 4.

Coons, E., & Kraehenbuehl, D. (1958). Information as a measure of structure in music. *Journal of Music Theory*, *2*(2), 127–161.

Cope, D. (2004). *Virtual music: computer synthesis of musical style*. MIT press.

Cope, D. (2005). *Computer models of musical creativity*. MIT Press Cambridge.

Crozier, W. R. (1999). The meanings of colour: Preferences among hues. *Pigment & Resin Technology*, *28*(1), 6–14.

Dahlstedt, P. (2004). *Sounds unheard of* (Unpublished doctoral dissertation). Chalmers Tekniska Högskola.

Danko-McGhee, K. (1999). Talismans of the carpatho-rusyn woman: The ritual practices and symbol system in the art of pysanky. *Journal of Art & Design Education*, *18*(3), 307–315.

Datta, S. (2010). Infinite sequences in the constructive geometry of tenth-century Hindu temple superstructures. In R. Kirkbride (Ed.), *Nexus network journal* (pp. 471–483). Birkhäuser Basel.

Davis, R. (1936). An evaluation and test of Birkhoff's aesthetic measure formula. *The Journal of General Psychology*, *15*(2), 231–240.

Davis, T., & Karamanlis, O. (2007). Gestural control of sonic swarms: composing with grouped sound objects. In *4th sound and music computing conference.*

Dawkins, R. (1986). *The blind watchmaker: Why the evidence of evolution reveals a universe without design*. WW Norton & Company.

DeCarlo, D., & Santella, A. (2002). *Stylization and abstraction of photographs* (Vol. 21). ACM.

Deng, X., Hui, S. K., & Hutchinson, J. W. (2010, October). Consumer preferences for color combinations: An empirical analysis of similarity-based color relationships. *Journal of Consumer Psychology*, *20*(4), 476–484.

den Heijer, E., & Eiben, A. E. (2010, April). Comparing aesthetic measures for evolutionary art. In *Evocomnet'10: Proceedings of the 2010 international conference on applications of evolutionary computation.* Springer-Verlag.

Di Scipio, A. (2003). 'Sound is the interface': from interactive to ecosystemic signal processing. *Organised Sound*, *8*(03), 269–277.

Doornbusch, P. (2002). A brief survey of mapping in algorithmic composition. In *Proceedings of the international computer music conference.*

Dorin, A. (2008). A survey of virtual ecosystems in generative electronic art. In *The art of artificial evolution* (pp. 289–309). Springer.

Drake, R. A. (1987). Effects of gaze manipulation on aesthetic judgments: Hemisphere priming of affect. *Acta Psychologica*, *65*(2), 91–99.

Draves, S. (2008). Evolution and collective intelligence of the electric sheep. In *The art of artificial evolution* (pp. 63–78). Springer.

Dubberly, H., Pangaro, P., & Haque, U. (2009). What is interaction? are there different types? *Interactions*, *16*(1), 69–75.

Dubnov, S. (2006). Spectral anticipations. *Computer Music Journal*, *30*(2), 63–83.

Duchowski, A. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods*, *34*(4), 455–470.

Duchowski, A. (2007). *Eye tracking methodology: Theory and practice* (Vol. 373). Springer London.

Ebcioğlu, K. (1986). An expert system for chorale harmonization. In *Proceeding for AAAI-86* (pp. 784–788).

Ebcioğlu, K. (1990). An expert system for harmonizing chorales in the style of JS Bach. *The Journal of Logic Programming*, *8*(1), 145–185.

Eglash, R. (1999). *African fractals: Modern computing and indigenous design*. Rutgers University Press.

Eiben, A. (2008). Evolutionary reproduction of Dutch masters: The Mondriaan and Escher evolvers. In *The art of artificial evolution* (pp. 211–224). Springer.

Eigenfeldt, A., & Pasquier, P. (2013). Evolving structures for electronic dance music. In *Proceedings of the 15th annual conference on genetic and evolutionary computation* (pp. 319–326).

Ellis, L., & Ficek, C. (2001). Color preferences according to gender and sexual orientation.

Personality and Individual Differences.

El-Said, I., El-Bouri, T., & Critchlow, K. (1993). *Islamic art and architecture: the system of geometric design*. Garnet Publishing.

Engelbrecht, A. P. (2006). *Fundamentals of computational swarm intelligence*. John Wiley & Sons.

Essl, K. (2007). Algorithmic composition. In N. Collins & J. d'Escrivan (Eds.), *The Cambridge companion to electronic music* (pp. 107–125). Cambridge University Press.

Eysenck, H. J. (1941). A critical and experimental study of colour preferences. *The American Journal of Psychology*, 385–394.

Fant, A. (1986). The case of the artist Hilma Af Klint. In *The spiritual in art: Abstract painting, 1890–1985* (pp. 155–163).

Fechner, G. T. (1876). *Vorschule der aesthetik* (Vol. 1). Breitkopf & Härtel.

Fernaeus, Y., Jonsson, M., & Tholander, J. (2012). Revisiting the Jacquard Loom: Threads of history and current patterns in HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1593–1602). ACM.

Fernández, J. D., & Vico, F. (2013). AI methods in algorithmic composition: a comprehensive survey. *Journal of Artificial Intelligence Research*, *48*(1).

Franke, H. W. (1971). Computers and visual art. *Leonardo*, *4*(4), 331–338.

Frazer, J. (1995). *An evolutionary architecture*. Architectural Association.

Fux, J. J., Mann, A., & Edmunds, J. (1965). *The study of counterpoint: from Johann Joseph Fux's 'Gradus ad Parnassum'*. W. W. Norton.

Galanter, P. (2003). What is generative art? complexity theory as a context for art theory. In *Proceedings of the 6th generative art conference* (pp. 271–282).

Galanter, P. (2012). Computational aesthetic evaluation: past and future. In *Computers and creativity* (pp. 255–293). Springer.

Garcia, R. (2001). Growing sound synthesizers using evolutionary methods. In *Proceedings ALMMA 2001: Artificial life models for musical applications workshop, ECAL 2001*.

Gärdenfors, P. (2004a). Conceptual spaces as a framework for knowledge representation. *Mind and Matter*, *2*(2), 9–27.

Gärdenfors, P. (2004b). *Conceptual spaces: The geometry of thought*. MIT Press.

Gell-Mann, M., & Lloyd, S. (1996). Information measures, effective complexity, and total information. *Complexity*, *2*(1), 44–52.

Gocer, A. (1999). A hypothesis concerning the character of Islamic art. *Journal of the History of Ideas*, *60*(4), 683–692.

Graham, G., et al. (2005). *Philosophy of the arts: An introduction to aesthetics*. Routledge.

Granger, G. (1955). An experimental study of colour preferences. *The Journal of General Psychology*, *52*(1), 3–20.

Green, O. (2008). Pondering value in the performance ecosystem. *Temps réel, improvisation et interactivité en électroacoustique/live-electronics—improvisation—interactivity in electroacoustics. eContact*, *10*.

Greenfield, G. R. (2008). Co-evolutionary methods in evolutionary art. In *The art of artificial evolution* (pp. 357–380). Springer.

Guilford, J., & Smith, P. C. (1959). A system of color-preferences. *The American Journal of Psychology*, *72*(4), 487–502.

Haji Mirza, S., Proulx, M., & Izquierdo, E. (2012). Reading Users' Minds from Their Eyes: A Method for Implicit Image Annotation. *Multimedia, IEEE Transactions on*(99), 1.

Hard, A., & Sivik, L. (2001). A theory of colors in combination—a descriptive model related to the ncs color-order system. *Color Research & Application*, *26*(1), 4–28.

Harder, H. (2012). Music on a long thin wire. *Leonardo Music Journal*, *22*, 87–87.

Hardoon, D., & Pasupa, K. (2010). Image ranking with implicit feedback from eye movements. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, 291–298.

Hemberg, M., O'Reilly, U.-M., Menges, A., Jonas, K., da Costa Gonçalves, M., & Fuchs, S. R. (2008). Genr8: Architects' experience with an emergent design tool. In *The art of artificial evolution* (pp. 167–188). Springer.

Hensel, M., Menges, A., & Weinstock, M. (Eds.). (2004). *Emergence: Morphogenetic Design Strategies (Architectural Design)*. Academy Press.

Heylighen, F. (2001). The science of self-organization and adaptivity. *The encyclopedia of life support systems*, *5*(3), 253–280.

Hild, H., Feulner, J., & Menzel, W. (1992). Harmonet: A neural net for harmonizing chorales in the style of JS Bach. In *Advances in neural information processing systems* (pp. 267–274).

Hiller, L. A., & Isaacson, L. M. (1959). *Experimental music; composition with an electronic computer*. Greenwood Publishing Group Inc.

Holmes, T., & Zanker, J. M. (2012). Using an oculomotor signature as an indicator of aesthetic preference. *i-Perception*, *3*(7), 426–439.

Hoover, A. K., Rosario, M. P., & Stanley, K. O. (2008). Scaffolding for interactively evolving novel drum tracks for existing songs. In *Applications of evolutionary computing* (pp. 412–422). Springer.

Horner, A., Beauchamp, J., & Packard, N. (1993). Timbre breeding. In *Proceedings of the 1993 international computer music conference.*

Hornof, A. J., & Cavender, A. (2005). EyeDraw. In *Proceedings of the 2005 SIGCHI conference on human factors in computing systems* (pp. 161–170). ACM.

Hume, D. (1985). Of the standard of taste. In E. Miller (Ed.), *Essays: Moral, political and literary.* Indianapolis: Liberty.

Itten, J. (1961). *The art of color; the subjective experience and objective rationale of color*. Wiley.

Itti, L. (2001). Computational modelling of visual attention. *Nature reviews neuroscience*, *2*(3).

Jacob, C., & Hushlak, G. (2008). Evolutionary and swarm design in science, art, and music. In *The art of artificial evolution* (pp. 145–166). Springer.

Jacob, R. J. K. (1990). What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the 1990 SIGCHI conference on human factors in computing systemssigchi conference* (pp. 11–18). ACM.

Jensen, K. (1999). *Timbre models of musical sounds* (Unpublished doctoral dissertation). Department of Computer Science, University of Copenhagen.

Jowers, I., Prats, M., EISSA, H., & LEE, J.-H. (2010). A study of emergence in the generation of Islamic geometric patterns. In *New frontiers: Proceedings of the 15th international conference on computer-aided architectural design research in asia* (pp. 7–10).

Jowers, I., Prats, M., McKay, A., & Garner, S. (2011). Design exploration with useless rules and eye tracking. In *Proceedings of the 18th international conference on engineering design.*

Jowers, I., Prats, M., McKay, A., & Garner, S. (2013). Evaluating an eye tracking interface for a two-dimensional sketch editor. *Computer-Aided Design*, *45*(5), 923–936.

Kaplan, M., & Cohen, E. (2003). Computer generated celtic design. *Rendering Techniques*, *3*, 9–19.

Karjalainen, M., & Smith, J. O. (1996). Body modeling techniques for string instrument synthesis. In *Proceedings of international computer music conference.*

Kirke, A., & Miranda, E. (2015). A multi-agent emotional society whose melodies represent its emergent social hierarchy and are generated by agent communications. *Journal of Artificial Societies and Social Simulation*, *18*(2), 16.

Klütsch, C. (2007). Computer graphic–aesthetic experiments between two cultures. *Leonardo*, *40*(5), 421–425.

Klütsch, C. (2012). Information aesthetics and the stuttgart school. In H. Higgins & D. Kahn (Eds.), *Mainframe experimentalism: Early computing and the foundation of the digital arts* (pp. 65–89). University of California Press.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, *4*, 219–227.

Kosorukoff, A. (2001). Human based genetic algorithm. In *Proceedings of the IEEE international conference on systems, man, and cybernetics* (pp. 3464–3469).

Krupinski, E., Locher, P., Nodine, C., & Mello-Thoms, C. (2007). Visual interest in pictorial art during an aesthetic experience. *Spatial Vision*, *21*(1), 55–77.

Kuehni, R. G. (2002). The early development of the munsell system. *Color Research & Application*, *27*(1), 20–27.

Langton, C. G. (1989). *Artificial life*. Addison-Wesley.

Langton, C. G. (1990). Computation at the edge of chaos: phase transitions and emergent computation. *Physica D: Nonlinear Phenomena*, *42*(1–3), 12–37.

Levine, M. W. (2000). *Fundamentals of sensation and perception*. Oxford University Press.

Lewis, G. E. (2000). Too many notes: Computers, complexity and culture in voyager. *Leonardo Music Journal*, *10*, 33–39.

Lewis, H. M., & Laland, K. N. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1599), 2171–2180.

Lewis, M. (2008). Evolutionary visual art and design. In *The art of artificial evolution* (pp. 3–37). Springer.

LeWitt, S. (1967). Paragraphs on conceptual art. *Artforum*, *5*(10), 79–83.

Li, M., & Vitányi, P. (2013). *An introduction to Kolmogorov complexity and its applications*. Springer Science & Business Media.

Lippmann, R. P. (1987). An introduction to computing with neural nets. *ASSP Magazine, IEEE*, *4*(2), 4–22.

Locher, J., Bool, F., & Ernst, B. (1992). *MC Escher, his life and complete graphic work*. Harry N. Abrams, Inc.

Loschky, L. C., & McConkie, G. W. (2000). User performance with gaze contingent multiresolutional displays. In *Proceedings of the ACM symposium on eye tracking research and applications* (pp. 97–103).

Loy, G. (1989). Composing with computers: a survey of some compositional formalisms and music programming languages. In M. Mathews & J. Pierce (Eds.), *Current directions in computer music research* (pp. 291–396). MIT Press.

MacCallum, R. M., Mauch, M., Burt, A., & Leroi, A. M. (2012). Evolution of music by public choice. *Proceedings of the National Academy of Sciences*, *109*(30), 12081–12086.

Machado, P., Romero, J., & Manaris, B. (2008). Experiments in computational aesthetics. In J. Romero & P. Machado (Eds.), *The art of artificial evolution: A handbook on evolutionary art and music* (pp. 381–415). Springer.

Machado, P., Romero, J., Santos, A., Cardoso, A., & Pazos, A. (2007). On the development of evolutionary artificial artists. *Computers & Graphics*, *31*(6), 818–826.

Machado, P., Romero, J., Santos, M. L., Cardoso, A., & Manaris, B. (2004). Adaptive critics for evolutionary artists. In *Applications of evolutionary computing* (pp. 437–446). Springer.

Majaranta, P., & Räihä, K.-J. (2002). Twenty years of eye typing. In *Proceedings of the 2002 symposium on eye tracking research and applications.*

Manaris, B., Machado, P., McCauley, C., Romero, J., & Krehbiel, D. (2005). Developing fitness functions for pleasant music: Zipf's law and interactive evolution systems. In *Applications of evolutionary computing* (pp. 498–507).

Manaris, B., Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W., & Davis, R. B. (2005). Zipf's law, music classification, and aesthetics. *Computer Music Journal*, *29*(1), 55–69.

Manaris, B., Roos, P., Machado, P., Krehbiel, D., Pellicoro, L., & Romero, J. (2007). A corpus-based hybrid approach to music analysis and composition. In *Proceedings of the national conference on artificial intelligence* (Vol. 22, p. 839).

Manousakis, S. (2006). Musical L-systems. *Unpublished master thesis, Institute of Sonology, The Hague*.

Markowsky, G. (1992). Misconceptions about the golden ratio. *The College Mathematics Journal*, *23*(1), 2–19.

Marks, L. (2006). Infinity and accident: Strategies of enfoldment in Islamic art and computer art. *Leonardo*, *39*(1), 37–43.

Matisse, H., & Flam, J. D. (1995). *Matisse on art*. University of California Press.

McCormack, J. (2003). Art and the mirror of nature. *Digital Creativity*, *14*(1), 3–22.

McCormack, J. (2005). Open problems in evolutionary music and art. *Applications of Evolutionary Computing*, 428–436.

McCormack, J., & Dorin, A. (2001). Art, emergence and the computational sublime. In *Proceedings of the second international conference on generative systems in the electronic arts* (pp. 67–81).

McDonald, R., & Smith, K. J. (2008). CIE94-a new colour-difference formula*. *Journal of the Society of Dyers and Colourists*, *111*(12), 376–379.

McHale, K. (2011). *Ada Lovelace: The unauthorized guide to the history, life and accomplishments of the first computer programmer*. Webster's Digital Services.

McMahon, J. A. (2011). Critical aesthetic realism. *The Journal of Aesthetic Education*, *45*(2).

McManus, I. C., Jones, A. L., & Cottrell, J. (1981). The aesthetics of colour. *Perception*, *10*(6), 651-666.

McWhinnie, H. J. (1968). A review of research on aesthetic measure. *Acta Psychologica*, *28*, 363–375.

Meyer, J. (2004). *Minimalism: art and polemics in the sixties*. Yale University Press.

Meyer, L. B. (1967). *Music, the arts and ideas: Patterns and predictions in twentieth-century culture*. University of Chicago Press.

Miranda, E. R. (1995). Granular synthesis of sounds by means of a cellular automaton. *Leonardo*, *28*(4), 297–300.

Miranda, E. R. (2002). Emergent sound repertoires in virtual societies. *Computer Music Journal*, *26*(2), 77–90.

Miranda, E. R. (2004). At the crossroads of evolutionary computation and music: Self-programming synthesizers, swarm orchestras and the origins of melody. *Evolutionary Computation*, *12*(2), 137–158.

Miranda, E. R., & Biles, A. (2007). *Evolutionary computer music*. Springer.

Miranda, E. R., & Junior, A. M. (2005). Granular synthesis of sounds through markov chains with fuzzy control. In *Proceedings of the international computer music conference* (pp. 193–196).

Mitchell, T. J., & Pipe, A. G. (2005). Convergence synthesis of dynamic frequency modulation tones using an evolution strategy. In *Applications of evolutionary computing* (pp. 533–538). Springer.

Moles, A. (1966). *Information theory and esthetic perception* (J. E. Coen, Trans.). University of Illinois Press.

Moon, P., & Spencer, D. E. (1944). Aesthetic measure applied to color harmony. *JOSA*, *34*(4), 234–242.

Moura, L., & Ramos, V. (2002). Swarm paintings—nonhuman art. *Architopia: Book, Art, Architecture, and Science, Institut d'Art Contemporain, Lyon/Villeurbanne, France*, 5–24.

Munsell, A. H. (1912). A pigment color system and notation. *The American Journal of Psychology*, 236–244.

Nake, F. (2005). Computer art: a personal recollection. In *Proceedings of the 5th conference on creativity & cognition.*

Narmour, E. (1977). *Beyond schenkerism*. University of Chicago Press.

Newman, M. E. (2005). Power laws, pareto distributions and zipf's law. *Contemporary physics*, *46*(5), 323–351.

Newton, I. (1979). *Opticks, or, a treatise of the reflections, refractions, inflections & colours of light*. Courier Corporation.

Nickerson, J., Sakamoto, Y., & Yu, L. (2011). Structures for creativity: The crowdsourcing of design. In *Proceedings of CHI workshop on crowdsourcing and human computation* (pp. 1–4).

Nielsen, M., & Stokes, M. (1998). The creation of the sRGB ICC profile. In *Prooceedings of the 1998 color and imaging conference* (pp. 253–257).

Nyman, M. (1999). *Experimental Music*. Cambridge University Press.

Ostwald, W. (1969). *The color primer*. Van Nostrand Reinhold.

Ou, L. C., & Luo, M. R. (2006). A study of colour harmony for two-colour combinations. *Color Reserch & Application*, *31*(3), 1–7.

Ou, L.-C., Luo, M. R., Woodcock, A. e., & Wright, A. (2004). A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *Color Research & Application*, *29*(3), 232–240.

Özdural, A. (2000). Mathematics and arts: Connections between theory and practice in the medieval islamic world. *Historia Mathematica*, *27*(2), 171–201.

Pallez, D., Collard, P., Baccino, T., & Dumercy, L. (2007). Eye-tracking evolutionary algorithm to minimize user fatigue in IEC applied to interactive one-max problem. In *the 2007 GECCO conference companion* (p. 2883).

Palmer, S. E., & Schloss, K. B. (2010). An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences*, *107*(19), 8877–8882.

Papadopoulos, G., & Wiggins, G. (1998). A genetic algorithm for the generation of jazz melodies. *Proceedings of STEP*, *98*.

Pask, G. (1961). *An approach to cybernetics*. Harper.

Pearce, M. T. (2005). *The construction and evaluation of statistical models of melodic structure in music perception and composition* (Unpublished doctoral dissertation). Department of Computing, City University, London.

Pearce, M. T., & Wiggins, G. A. (2012). Auditory expectation: the information dynamics of music perception and cognition. *Topics in cognitive science*, *4*(4), 625–652.

Perlin, K., & Hoffert, E. M. (1989). Hypertexture. In *ACM SIGGRAPH computer graphics* (Vol. 23, pp. 253–262). ACM.

Perrett, D., Lee, K., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D., ... Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, *394*(6696), 884–887.

Phon-Amnuaisuk, S., & Wiggins, G. (1999). The four-part harmonisation problem: a comparison between genetic algorithms and a rule-based system. In *Proceedings of the aisb'99 symposium on musical creativity* (pp. 28–34).

Pieters, R., & Warlop, L. (1999). Visual attention during brand choice: The impact of time pressure and task motivation. *International Journal of Research in Marketing*, *16*(1), 1–16.

Plumhoff, J. E., & Schirillo, J. A. (2009). Mondrian, eye movements, and the oblique effect. *Perception*, *38*(5), 719–731.

Ponsford, D., Wiggins, G. A., & Mellish, C. S. (1999). Statistical learning of harmonic movement. *Journal of New Music Research*, *28*(2), 150-177.

Potter, K., Wiggins, G. A., & Pearce, M. T. (2007). Towards greater objectivity in music theory: Information-dynamic analysis of minimalist music. *Musicae Scientiae*, *11*(2), 295–324.

Prats, M., Earl, C., Garner, S., & Jowers, I. (2006). Shape exploration of designs in a style: Toward generation of product designs. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *20*(3), 201–215.

Prusinkiewicz, P., & Lindenmayer, A. (1996). *The Algorithmic Beauty of Plants*. Springer.

Quinn, A. J., & Bederson, B. B. (2011). Human computation: a survey and taxonomy of a growing field. In *Proceedings of the 2011 SIGCHI conference on human factors in computing systems* (p. 1403-1412). ACM.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422.

Reed, K. (2013). Aesthetic measures for evolutionary vase design. In P. Machado, J. McDermott, & A. Carballal (Eds.), *Lecture notes in computer science* (pp. 59–71). Springer.

Reich, S. (2004). *Writings on music, 1965-2000* (P. Hillier, Ed.). Oxford University Press.

Reichardt, J. (1970). *Cybernetic serendipity, the computer and the arts.* Studio International Special Issue. London.

Reingold, E. M., Loschky, L. C., McConkie, G. W., & Stampe, D. M. (2003). Gaze-contingent multiresolutional displays: An integrative review. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(2), 307–328.

Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. In *ACM SIGGRAPH computer graphics* (Vol. 21, pp. 25–34). ACM.

Rigau, J., Feixas, M., & Sbert, M. (2007). Conceptualizing birkhoff's aesthetic measure using shannon entropy and kolmogorov complexity. *Computational Aesthetics*, 1–8.

Roads, C. (2004). *Microsound*. The MIT Press.

Romero, J., Machado, P., Santos, A., & Cardoso, A. (2003). On the development of critics in evolutionary computation artists. In *Applications of evolutionary computing* (pp. 559–569). Springer.

Romero, J. J., & Machado, P. (2007). *The art of artificial evolution: a handbook on evolutionary art and music*. Springer Science & Business Media.

Rothgeb, J. (1980). Simulating musical skills by digital computer. *Computer Music Journal*, *4*(2), 36–40.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52.

Santella, A., & DeCarlo, D. (2004). Robust clustering of eye movement recordings for quantification of visual interest. In *Proceedings of the 2004 symposium on eye tracking research and applications* (pp. 27–34).

Scherffig, L. (2005). It's in your eyes. *ZKM Institute for Basic Research, Karlsruhe*.

Schillinger, J. (1948). *The mathematical basis of the arts*. Philosphical Library.

Schloss, K. B., & Palmer, S. E. (2011). Aesthetic response to color combinations: preference, harmony, and similarity. *Attention, Perception, & Psychophysics*, *73*(2), 551–571.

Schloss, K. B., Poggesi, R. M., & Palmer, S. E. (2011). Effects of university affiliation and "school spirit" on color preferences: Berkeley versus stanford. *Psychonomic bulletin & review*, *18*(3), 498–504.

Schnier, T., & Gero, J. (1998). From Mondrian to Frank Lloyd Wright: Transforming evolving representations. In I. Parmee (Ed.), *Adaptive computing in design and manufacture* (p. 207-219). Springer London.

Schon, D. A., & Wiggins, G. (1992). Kinds of seeing and their functions in designing. *Design studies*, *13*(2), 135–156.

Serexhe, B. (Ed.). (2013). *Preservation of digital art: Theory and practice*. Ambra V.

Shalem, A. (2012). What do we mean when we say "Islamic Art"? An urgent plea for a critical re-writing of the history of the arts of Islam. *Journal of Art Historiography*, *6*.

Shannon, C. E., & Weaver, W. (1964). *The Mathematical Theory of Communication*. University of Illinois Press.

Shevell, S. K. (2003). *The science of color*. Elsevier.

Shimamura, A. P. (2012). Toward a science of aesthetics. In A. P. Shimamura & S. E. Palmers (Eds.), *Aesthetic science: Connecting minds, brains, and experience* (pp. 3–28). Oxford University Press.

Sims, K. (1991). Artificial evolution for computer graphics. *Computer Graphics*, *25*(4), 319–328.

Smith, B. (1988). *Foundations of gestalt theory*. Philosophia Verlag Gmbh.

Smucker, J., Shaw, R., & Cunningham, J. (2009). *Amish abstractions: Quilts from the collection* *of faith and stephen brown*. Pomegranate Communications.

Solomos, M. (2006). Cellular automata in Xenakis's music. Theory and practice. In *Proceedings of the international symposium Iannis Xenakis* (pp. 11–15).

Starker, I., & Bolt, R. A. (1990). A gaze-responsive self-disclosing display. In *Proceedings of the 1990 SIGCHI conference on human factors in computing systems* (pp. 3–10). ACM.

Stellmach, S., & Dachselt, R. (2012). Look & touch: gaze-supported target acquisition. In *Proceedings of the 2012 SIGCHI conference on human factors in computing systems.*

Stiny, G., & Gips, J. (1972). Shape grammars and the generative specification of painting and sculpture. *Information processing*, *71*, 1460–1465.

Strobl, G., Eckel, G., Rocchesso, D., & le Grazie, S. (2006). Sound texture modeling: A survey. In *Proceedings of the 2006 sound and music computing (smc) international conference* (pp. 61–5).

Surowiecki, J. (2005). *The wisdom of crowds*. Anchor.

Suzuki, Y., & Takeshima, H. (2004). Equal-loudness-level contours for pure tones. *The Journal of the Acoustical Society of America*, *116*(2), 918.

Tanriverdi, V., & Jacob, R. J. K. (2000). Interacting with eye movements in virtual environments. In *Proceedings of the 2000 SIGCHI conference on human factors in computing systems* (pp. 265–272). ACM.

Todd, P., & Loy, G. (2003). Connectionism and musiconomy. In *Music and connectionism* (pp. 20–36). MIT Press.

Todd, P. M., & Werner, G. M. (1999). Frankensteinian methods for evolutionary music. In *Muiscal networks* (pp. 313–339). ACM.

Todd, S., & Latham, W. (1994). *Evolutionary art and computers*. Academic Press, Inc.

Toiviainen, P. (1995). Modeling the target-note technique of bebop-style jazz improvisation: An artificial neural network approach. *Music Perception*, *12*(4), 399–413.

Trivedi, K. (1989). Hindu temples: Models of a fractal universe. *The Visual Computer*, *5*(4), 243-258.

Tuisku, O., Majaranta, P., Isokoski, P., & Räihä, K.-J. (2008). Now Dasher! Dash away!: longitudinal study of fast text entry by eye gaze. In *Proceedings of the 2008 symposium on eye tracking research and applications* (pp. 19–26).

Turing, A. M. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the*

Royal Society of London B: Biological Sciences, *237*(641), 37–72.

Vaggione, H. (2001). Some ontological remarks about music composition processes. *Computer Music Journal*, *25*(1).

Von Foerster, H. (2003a). Disorder/order: discovery or invention? In *Understanding understanding* (pp. 273–282). Springer Science & Business Media.

Von Foerster, H. (2003b). On self-organizing systems and their environments. In *Understanding understanding* (pp. 1–19). Springer Science & Business Media.

Von Foerster, H. (2007). *Understanding understanding: Essays on cybernetics and cognition*. Springer Science & Business Media.

Von Goethe, J. W., & Eastlake, C. L. (1970). *Theory of colours*. MIT Press.

Von Neumann, J., Burks, A. W., et al. (1966). Theory of self-reproducing automata. *IEEE Transactions on Neural Networks*, *5*(1), 3–14.

Waschka II, R. (2007). Composing with genetic algorithms: Gendash. In *Evolutionary computer music* (pp. 117–136). Springer.

Waters, S. (2007). Performance ecosystems: Ecological approaches to musical interaction. *EMS: Electroacoustic Music Studies Network*.

Westland, S., Laycock, K., Cheung, V., Henry, P., & Mahyar, F. (2007). Colour harmony. *Journal of the International Colour Association*(1), 1–15.

Whitelaw, M. (2004). *Metacreation: art and artificial life*. MIT Press.

Wiener, N. (1950). Cybernetics. *Bulletin of the American Academy of Arts and Sciences*, *3*(7), 2–4.

Wiener, N. (1961). *Cybernetics or control and communication in the animal and the machine*. MIT press.

Wishart, T., & Emmerson, S. (1996). *On Sonic Art*. Psychology Press.

Wooding, D. S. (2002, November). Eye movements of large populations: II. Deriving regions of interest, coverage, and similarity using fixation maps. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 518–528.

Worth, P., & Stepney, S. (2005, March). Growing music: musical interpretations of l-systems. In *Ec'05: Proceedings of the 3rd european conference on applications of evolutionary computing.* Springer-Verlag.

Wundt, W. (1897). *Outlines of psychology*. Englemann.

Xenakis, I. (1992). *Formalized Music*. Pendragon Press.

Yarbus, A. L. (1967). *Eye movements during perception of complex objects*. Springer.

Youngblood, J. E. (1958). Style as information. *Journal of Music Theory*, *2*, 24–35.

Yu, H., & Winkler, S. (2013). Image complexity and spatial information. In *Proceedings of the 2013 international workshop on quality of multimedia experience (QoMEX)* (pp. 12–17). IEEE.

Yu, L., & Sakamoto, Y. (2011, July). Feature selection in crowd creativity. In *Fac'11: Proceedings of the 6th international conference on foundations of augmented cognition: directing the future of adaptive systems.* Springer-Verlag.

Zhai, S., Morimoto, C., & Ihde, S. (1999). Manual and gaze input cascaded (MAGIC) pointing. In *Proceedings of the 1999 SIGCHI conference on human factors in computing systems* (pp. 246–253). ACM.